

BESZÉDKUTATÁS 1996

TANULMÁNYKÖTET



MTA Nyelvtudományi Intézete
Budapest

BESZÉDKUTATÁS 96

Szerkesztette:

Gósy Mária

A Magyar Tudományos Akadémia Nyelvtudományi Intézete
Budapest 1996

Az egyes tanulmányokat lektorálták:

Gósy Mária

Olaszy Gábor

ISBN 963 9074 01 2

© MTA Nyelvtudományi Intézete

Felelős kiadó: Kiefer Ferenc igazgató

Tartalom

Gósy Mária: A magánhangzók minősége a spontán beszéd szavaiban	1
Nikléczy Péter: Beszélő személy azonosítása szűk frekvenciás szavak alapján	20
Kovács Andrea: Fonetikai jelenségek egy kétéves gyermek nyelvében	31
Gósy Mária: A transzformációs észlelés fejlődése és zavarai	44
Balázs Boglárka–Gósy Mária–Szabó Iván: A gége nélküli beszéd fonetikai sajátosságai	58
Olaszy Gábor: Számok kiejtésének fonetikai vizsgálata	72
Szalai Enikő: A lexikális hozzáférés ép és zavart folyamatai	79
Olaszy Gábor: Szabályrendszer prozódiai elemek gépi megvalósításához	97

A MAGÁNHANGZÓK MINŐSÉGE A SPONTÁN BESZÉD SZAVAIBAN

Gósy Mária
MTA Nyelvtudományi Intézete

Bevezetés

A magánhangzók akusztikai szerkezetének vizsgálata immár több évtizedes múltra tekinthet vissza. A metodológiai sajátosságok miatt az elemzések jobbára olvasott szövegrészeket, többnyire izolált szavak magánhangzóira vonatkoztak (Tarnóczy 1941; Molnár 1970; Bolla 1982). A beszéd szintézis megjelenésével lehetővé vált egyfajta ellenőrzés, amely az elemzési adatok helyességére, illetőleg korrekciójára vonatkozott (Olaszy 1985). A természetesen ejtett magánhangzók minőségének és az akusztikai szerkezetnek az összefüggésére is végeztek néhány kísérletet (Tarnóczy 1965; Gósy 1989; Szalai 1995). Nincsenek azonban megbízható adataink a spontán beszédben előforduló szavak magánhangzói minőségének és akusztikai szerkezetének kapcsolatáról. A magyar nyelv nem redukálja - nyelvspecifikus jelenséggént - a magánhangzókat, mint ahogyan ez megtörténik az angolban (Jones 1966) vagy a németben (Lindner 1981, 238). A magánhangzók redukálódása az említett nyelvekben azt jelenti, hogy egy adott hangminőség semleges magánhangzóként, svaként jelenik meg a nyelvi szabályoknak megfelelően, rendszerint mint egy hangsúlytalan szótag magánhangzó fonémájának realizációja. A sva tehát minden esetben valamely fonéma redukált változata, a hollandban például mindig rövidebb időtartamban jelentkezik, mint a rövid magánhangzók (Koopmans-van Beinum 1994). A svat mint központi magánhangzót említi Laziczius (vö. 1963, 73).

A spontán beszédben lévő szavak időviszonyai természetesen nem azonosak az izolált ejtésű szavakéival. kövekezőképpen az időtartambeli különbség a szavakat alkotó beszédhangokban jelentkezik. A folyamatos beszédben egyfajta természetes rövidülés jelentkezik a magyarban is, elsősorban a magánhangzók ejtésekor, ennek rendszeres elemzéséről azonban nincsenek megbízható adatok. Ez a fajta rövidülés elméletileg nem jelent egyúttal minőségbeli változást is; a magyar magánhangzók,

függetlenül a hangsorban betöltött helyzetüktől és a közvetlen kontextustól, mindig teljes értékűen képződnek, vagyis anyanyelvünkben minőségileg redukált magánhangzók mint bizonyos fonémák szabályszerűen megjelenő realizációi nincsenek. Eltekintünk most azoktól a sva-jelenségektől, amelyek rendszerint két mássalhangzó között, fiziológiai kényszerként vagy szóhatárvég-mutatóként jelentkeznek (vö. Kassai 1978). Ezekben az esetekben ui. egy valóban "semleges" magánhangzószerű hangrészlet iktatódik a folyamatos beszédbe különféle funkcióban.

A beszédtempó gyorsulása mindazonáltal felveti annak a lehetőségét, hogy a huszadik század végére vajon nem jelentek-e meg a spontán magyar beszéd szavaiban is a redukált magánhangzók, amelyek tehát hangminőségüket tekintve már nem vagy alig különböznek egymástól. A kérdés az, hogy például a "*Kitaláltatok az eredetit?*" mondat magánhangzói spontán beszédben miként jelenhetnek meg. Az alternatívák a következők:

- a) [kito la:ltɑ:tok əz eredetit]
- b) [kitɔla:ltɑ:tək əz eredetit]
- c) [kitɔla:ltətək əz erədətɪt]

Az első eset megfelel a nyelvspecifikus elvárásoknak; függetlenül az aktuális időtartamtól, a kiejtett magánhangzók minősége megfelel a tervezett magánhangzók minőségének. A második esetben részleges minőségi redukció jelenik meg; a nem hangsúlyos helyzetben lévő magánhangzók egy része minőségileg más hangba csap át, azaz a kiejtett magánhangzó nem felel meg a tervezett beszédhangnak. A harmadik esetben - teoretikusan - valamennyi lehetséges redukálható pozícióban lévő beszédhang elveszíti eredeti (tervezett) minőségét. A hipotézis az, hogy a hangsúlyos helyzetű magánhangzó minősége állandó, vagyis a realizált beszédhang mindenkor megfelel a tervezett beszédhangnak.

A spontán beszéd hallgatásakor - az átlagos beszédtempótól csaknem függetlenül - a beszélő/hallgatónak mindig az a benyomása, hogy teljes értékűen képzett magánhangzót hall. Az esetleges redukciók csak akkor tűnnek fel (és akkor sem minden esetben), ha azok egész szótagot érintenek vagy egy hang teljes kiesésével járnak. A magyarázat a beszédmegértési mechanizmus működésében rejlik. A beszédészlelési szintekkel kvázi-parallel módon működő lexikális hozzáférés folyamatában a

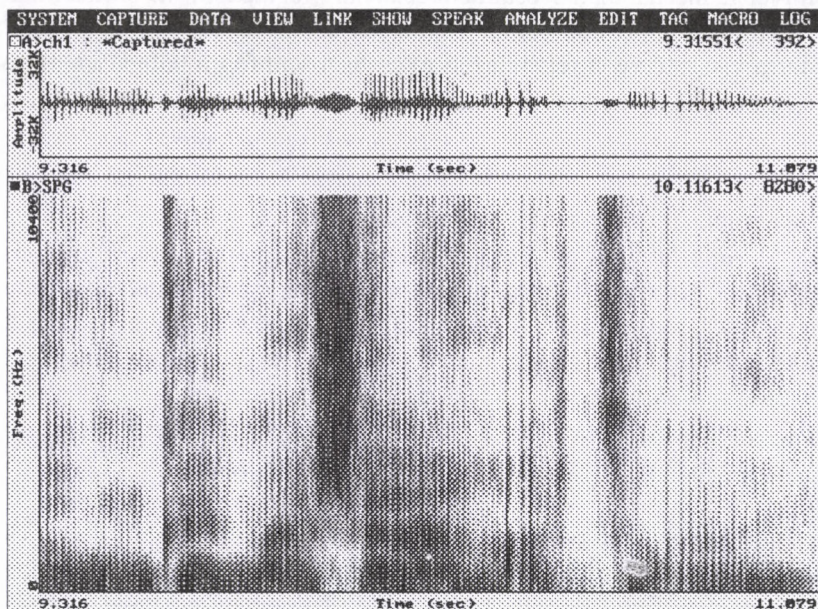
felsőbb szint(ek) felülbírálják az alsóbb szinteken hozott döntési eredményeket, s ez fogja eredményezni az adott szemantikának megfelelő hangsor felismerését (Gósy 1989). Normál kommunikációs szituációban a beszédmegértés számára meghatározó a kontextushatás, a beszédészlelés akusztikai és fonetikai szintjének működése nem ítéltető meg, mivel a mechanizmus csaknem azonnal "felkapcsol" a szemantikai és morfológiai/morfofonológiai elemzések szintjére. Az előbbi példamondatra visszatérve, a feltételezett c) esetben sem fog a hallgató a második és az utolsó szótagban [ɔ]-t észlelni, mivel a real-time működő lexikális hozzáférés eredményeként a *kitaláltatok* szó jelentését azonosítja, amelyben a tervezett [ɔ] és [o] magánhangzók fordulnak elő függetlenül az aktuális realizálásuktól az adott szótagokban.

A **hipotézisünk** az, hogy a beszédtempó gyorsulásának, valamint a spontán beszéd artikulációs sajátosságainak következtében a magánhangzók redukálódnak a magyarban is, de ez a redukció nem korlátozódik az időtartamra, hanem érinti a magánhangzó minőségét is. Hasonló jellegű vagy célú vizsgálatot a magánhangzók viselkedésére a folyamatos beszédben korábban nem végeztek. Kísérletet terveztünk a jelenség beható akusztikai és percepciós vizsgálatára.

Anyag és módszer

Különböző témájú spontán hanganyagot rögzítettünk egy magyar anyanyelvű, fiatal felnőtt férfi beszélővel (magyar anyanyelvű beszélő, hosszabb ideig nem élt külföldön). Az adatközlő átlagos artikulációs tempója 12,75 beszédhang/s volt; a szótástartomány pedig 11,2-14,3 beszédhang/s. Az átlagos beszédtempója (a beszédszünetek is figyelembe véve): 9,8 beszédhang/s. Ezeknek az értékeknek megfelelően kimondható, hogy a választott adatközlő beszédtempója közepes, artikulációs tempója is átlagos (se nem gyors, se nem lassú). A teljes hanganyag időtartama 20 perc, ebből a jelen vizsgálathoz mintegy 6 percnyi részt dolgoztunk fel. Az első vizsgálatsorozatban egyidejű auditív és vizuális visszacsatolás (és megerősítés) mellett szegmentáltuk a spontán beszédben lévő szavak valamennyi magánhangzóját, függetlenül a hangsorban elfoglalt helyétől, illetőleg a hangsúlyviszonyoktól. A szegmentálás során a magánhangzók ún. tiszta fázisának

kijelölésére törekedtünk, vagyis az átmeneti részeket elhagytuk (1. ábra). A szegmentálás után meghatároztuk a magánhangzók időtartamát (ms-ban) és az első három formáns értékét (Hz-ben). Az akusztikai elemzéseket a Kay 345-es digitális jelfeldolgozó különböző beállításával végeztük.



1. ábra

A magánhangzók szegmentálásának módja

A második kísérletsorozatban a szegmentált magánhangzók közül ötöt választottunk ki, és ezeket random sorrendben, kétszeri ismétléssel magnetofonszalagra rögzítettünk. A kiválasztott magánhangzók a következők: a velárisok közül az [a:, ɔ, o], a palatálisok közül az [e, ø, ø:]. A kiválasztás a hangminőség változására irányult. A redukált [ɔ] magánhangzóhoz legközelebbi [ø, ø:] hangok mintegy "etalonként" szerepeltek; az elemzett hanganyagban közel azonos arányban fordult elő a nyelvileg

rövid és hosszú labiális palatális hang. A szegmentált beszédhangok lehallgatása alapján azokat a magánhangzótípusokat választottuk ki, amelyek gyakran realizálódtak a semleges magánhangzóban. A vizsgált korpuszban a felső és középső nyelvállású illabiális palatális magánhangzók egyetlen egyszer sem realizálódtak svaként, hasonlóképpen a felső nyelvállású labiális veláris sem. A felső nyelvállású labiális palatális magánhangzó elemzésére - bár minősége nemegyszer a svaba csapott át - a korpuszban jelentkező rendkívül kicsiny elemszáma miatt nem volt mód. Ennek megfelelően vizsgáltuk a hanganyagban előforduló összes veláris legalsó, alsó és középső nyelvállású, valamint a palatális alsó és a palatális labiális középső nyelvállású magánhangzót (a hosszú [o:] mindössze két esetben fordult elő, ezeket az analízisben nem vettük figyelembe).

A percepció tesztelésre tehát a felsorolt ötféle magánhangzó előfordulásait készítettük elő. A lehallgatásra kész hanganyag összesen 710 magánhangzót tartalmazott a spontán beszédbeli eloszlásnak megfelelően, random sorrendben (1. táblázat). A tesztelést 50-es csoportonként végeztük, szünetekkel. Egy kísérletben 200 magánhangzót hallgattattunk le.

A kísérleti személyek egyetemi hallgatók voltak, átlagéletkoruk 21 év, összesen 10 fő (6 nő, 4 férfi). A feladatuk a hallott magánhangzó betűjelének leírása volt, azaz olyan módszert választottunk, amelyben a lehallgatók döntését a magyar magánhangzó-rendszerre korlátoztuk. Természetesen a 'nem azonosítható' kategória használatát lehetővé tettük, jelölés ebben azonban egyetlen kísérleti személy tesztlapján sem fordult elő. A tiszta fázisukat tekintve szegmentált magánhangzók tehát a magyar hangrendszerben egyértelműen azonosíthatók voltak. A feldolgozás során az akusztikai adatokat és a percepció adatokat vetettük össze.

A 710 magánhangzóra egyenként 4 akusztikai paramétert határoztunk meg, valamint a 10 kísérleti személy percepció adatát, így összesen 28400 adatot kaptunk a magánhangzókra. A palatális, középső nyelvállású nyelviileg rövid és hosszú magánhangzókat nem választottuk külön az elemzésekben. A statisztikai elemzések közül szignifikanciavizsgálatot és korrelációs számítást végeztünk.

1. táblázat: A spontán beszédből szegmentált magánhangzók száma

Magánhangzó (IPA-átírásban)	Előfordulás	
	száma	százaléka
a:	84	11,8
ɔ	246	34,6
o	120	16,9
ø, ø:	39	5,4
e	221	31,3
összesen	710	100

Eredmények

Kiindulásul meghatároztuk, hogy az [ø, ø:] magánhangzóknek melyek az elemzett akusztikai adatai a vizsgált beszélő ejtésében (2. táblázat). A tesztelési adatok szerint ezt a beszédhangot 100%-ban minden adatközlő helyesen azonosította. Az [ø]-re kapott átlagértékek a következők, időtartam: 87 ms, az F1: 410 Hz, az F2: 1550 Hz, az F3: 2320 Hz. A kísérleti személyek kivétel nélkül rövid [ø] magánhangzókat azonosítottak; a nyelvileg rövidek és hosszúak között alig van fizikai időtartam-különbség. A nyelvileg rövidek átlaga: 74,2 ms, a hosszúaké: 84,8 ms.

2. táblázat: Az [ø] magánhangzó időtartam- és formánsértékei

Időtartam (ms)	Formáns-határértékek (Hz)		
	F1	F2	F3
48-70	380-420	1440-1670	2000-2400
80-180	400-450	1350-1700	2260-2480
<hr/>			
48-180	380-450	1350-1700	2000-2480

A további elemzések során a percepció adatokat tekintettük irányadóknak. A 10 kísérleti személy percepció ítéletét akkor tekintettük az adott magánhangzóra érvényesnek, ha legalább 8 kísérleti személy azo-

nosan döntött (minimum 80%). (A 80%-nál alacsonyabb percepció az azonosság esetén az adott magánhangzót az elemzések során nem vettük figyelembe).

Az eredmények azt mutatják, hogy a vizsgált magánhangzókat a kísérleti személyek kétféleképpen azonosították: vagy felismerték az eredeti magánhangzót, vagy a középső nyelvállású, rövid palatális magánhangzót észlelték. Más hangminőség nem jelent meg az észleletekben (a teljes korpuszt tekintve mindössze 5 magánhangzó esetében volt 80%-nál kisebb percepció az azonosság). A 3. táblázatban összesítettük a 4 vizsgált magánhangzóra kapott azonosítási eredményeket. (Munkamegoldásként a percepció tesztlapokon *ő*-vel jelölt azonosításokat az elemzések során svakként kezeljük.)

3. táblázat: A vizsgált magánhangzók percepció adatai

Magánhangzó	Azonosítás (%)	
	eredetiként	svakként
a:	46,4	53,6
ɔ	46,2	53,8
o	67,5	32,5
e	42,6	57,4
átlag	50,6	49,3

A táblázat adatai azt mutatják, hogy - a kísérleti személyek egyöntetű felismerése alapján - a vizsgált magánhangzók mindegyike a spontán beszédben semleges hangminőségként jelenik meg. Szignifikáns különbséget látunk az [o] és a másik három beszédhang között, az utóbbiakhoz viszonyítva a középső nyelvállású veláris hang őrzi meg legnagyobb mértékben az eredeti hangminőséget. Az ok az artikulációban rejlik, tekintettel arra, hogy a veláris középső nyelvállású magánhangzó csupán egy jegyben tér el a palatális középső nyelvállásútól, a beszélő nyilván ösztönösen pontosabb képzésre törekszik a megkülönböztethetőség érdekében. A már 2 jegynyi eltérés esetén - az [ɔ] és az [ø] között a nyelv vízszintes és függőleges mozgását illetően, valamint az [c] és az [ø] között a nyelv függőleges mozgása és az ajakartikuláció

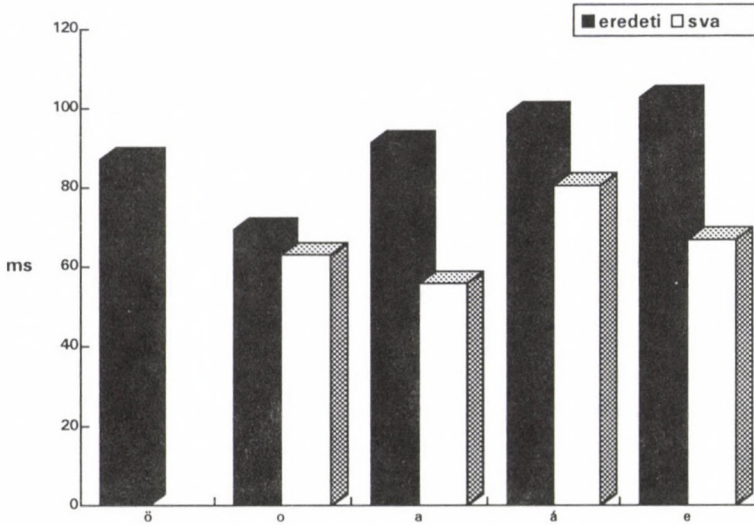
tekintetében - az artikulációs kivitelezés lényegesen pontatlanabb, kevésbé törekszik a megkülönböztetésre. Az [a:] és az [ø] közötti 3 jegynyi eltérés (a nyelv vízszintes és függőleges mozgása, valamint az ajakműködés) hasonló eredménnyel jár.

Az elemzések során a magánhangzók időtartamát és az első két formáns értékét vetettük össze a percepciós adatokkal. Minthogy a harmadik formánsnak a magyar magánhangzók azonosításában redundáns szerepe van, ezért a jelen elemzésekben az F3-ra nem tértünk ki. A 4. táblázatban a magánhangzók akusztikai szerkezetére kapott szórásértékeket adjuk meg.

4. táblázat: Az elemzett magánhangzók akusztikai adatai

eredeti magánhangzó			[ə]		
V idő	F1	F2	idő	F1	F2
o 30-90	340-510	720-1100	40-85	310-500	1075-1600
o 50-210	480-600	960-1230	25-155	390-5401	160-1500
a: 80-155	600-760	1335-1510	50-124	395-6201	415-1770
e 47-200	480-585	1460-1835	28-130	370-570	1060-1785

Az adatok számos releváns különbségről árulkodnak. Az időtartamok szórásstartománya kisebb, az értékek alacsonyabbak a semleges magánhangzó felismerésekor, vagyis a rövidebb időtartam a magyarban is jellemző a hangminőség módosulásakor. Valamennyi sva-ként azonosított hangot tekintve, a magyar redukált magánhangzók hasonló időtartam-határértékek között realizálódnak, mint például a hollandok. Az utóbbiakra a 18-120 ms-os értékhatárok a jellemzőek, a magyarra a 25-130 ms; a hollandban a gyakori időtartamérték a 26-70 ms, 47 ms-os átlagértékkel, a magyarban az átlagérték 62 ms. Nem mondható ugyanakkor ki, hogy a sva felismerése csak az időtartam függvénye, hiszen rövidebb időtartamú magánhangzókat is azonosítottak az eredetinek megfelelően a kísérleti személyek, míg hosszabb időtartamúakat is a semleges magánhangzóként (2. ábra).

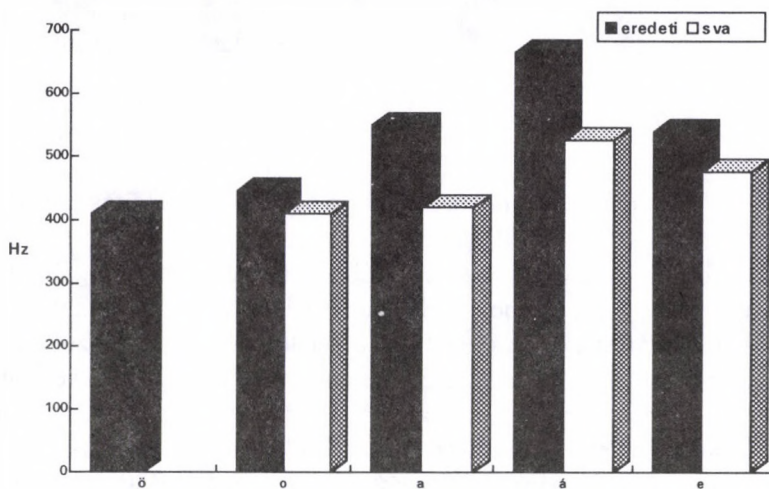


2. ábra

Az időtartamátlagok alakulása a kétféle azonosítás esetén

A formánsértékek elemzése további információval szolgált az azonosítások magyarázatához. Jellegetes különbséget mutatnak az első formáns határértékei az eredeti és a semleges magánhangzót tekintve. Lényegesen szűkebb a svaként azonosított hangok frekvenciatartománya, s alacsonyabban realizálódik. Ha összevetjük az összes semleges magánhangzóként azonosított F1-értéket és ugyanazon beszélő [ø]-ejtésének első formánsaira kapott adatokat ($\phi F1=380-450$ Hz), azt látjuk, hogy az eredeti [a:] magánhangzó sva-észleletének kivételével a határértékek nagyon hasonlóak: $\phi F1=310-500$ Hz, $\phi F1=390-540$ Hz és $eF1=370-570$ Hz. A második formánsok értékei az eredeti [ø] magánhangzónál: $F2=1350-1675$ Hz, ehhez képest a svaként azonosított hangoknál ez a tartomány lényegesen szélesebb, mind az alacsonyabb, mind a magasabb frekvenciák felé eltolódva. Ez egyértelműen arra utal, hogy az eredetileg ejtett magánhangzó artikulációs sajátosságai valamilyen mértékben a második formánsok értékében őrződnek. Az

adott beszélőre jellemző [ø]-akusztikai szerkezet kisebb-nagyobb mértékben azonban valamennyi sva-ként azonosított magánhangzóban megjelenik. A vizsgált korpuszban a svara kapott F1-határértékek: 313-622 Hz, az átlag 486 Hz, az F2 határértékei pedig 886-1785 Hz, az átlag 1457 Hz. A nem fonéma-realizációként jelentkező svara - hagyományos módszerrel - mért formánsértékekre a következő adatokat közölték: F1=210-660 Hz, F2=1050-2410 Hz (vö. Kassai 1978, 95). Összehasonlításképpen, a holland semleges magánhangzók F1-átlaga 335 Hz, az F2-é 1285 Hz (Koopmans-van Beinum i.m.); az angolé: F1=480-580 Hz, F2=1350-1530 Hz (Fry 1976, 93). Az eredetinek és a semlegesnek felismert magánhangzók F1-értékei anyagunkban szignifikánsan különböznek (3. ábra).



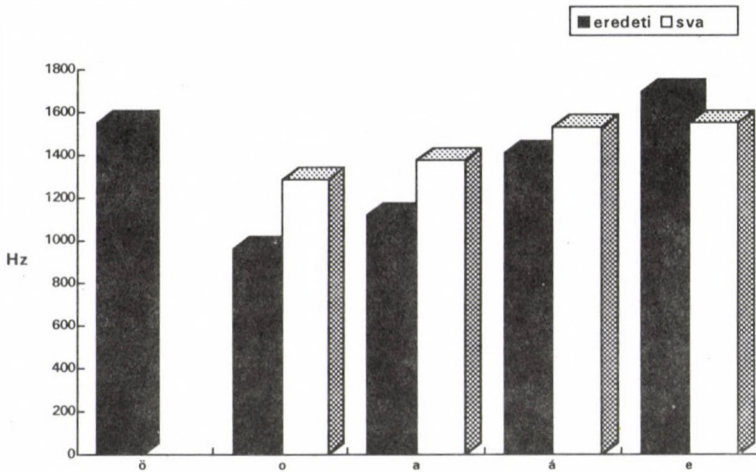
3. ábra

Az első formáns átlagértékei az azonosítás függvényében

A 4. ábrán az eredeti [ø]magánhangzónak és az elemzett többi magánhangzónak az F2-értékeit ábrázoltuk. Jól látható itt is, hogy a második formánsok átlagértékei a sva-azonosítás esetén közelítik az [ø]magánhangzó F2-jének átlagát, míg az eredeti magánhangzó pontos

azonosításakor ez nem áll. Jellegzetes ugyanakkor az is, hogy a svaként azonosított magánhangzó második formánsának átlagértékei az eredeti magánhangzóra jellemző értékeket is magukban hordozzák.

A korreláció-számítással arra igyekeztünk választ kapni, hogy a három elemzett változó (időtartam, F1 és F2) hogyan viszonylik az adott magánhangzó percepciójához. Melyik tényező esetében találunk szoros, hol kevésbé szoros összefüggést, illetőleg hogy esetleg az adott faktor nem meghatározó? Az 5. táblázatban a korrelációs számítás eredményeit foglaltuk össze, az értékeket 5 tizedesjegyre megadva (rövidítések: idő=fizikai időtartam, percep.=percepció, észlelet, r=korreláció értéke, korr.=korreláció).



4. ábra

A második formáns átlagértékei az azonosítás függvényében

Az időtartam és a magánhangzó minőségének azonosítása az [o]-t kivéve közepes korrelációt mutat, vagyis az időtartam befolyással van a percepcióra, azonban elsőrendűen nem meghatározó. Az [o] esetében ugyanakkor az időtartam egyáltalán nem korrelált a percepcióval. Az első formáns értéke és a felismert hangminőség megegyezik az [o] kivételével

- ahol a korreláció közepes - a három másik magánhangzó esetében erős összefüggést mutat. A második formáns értéke és a felismerés erős korrelációt mutat az [o] és az [a:], közepes korrelációt az [ɔ] és az [e] magánhangzók esetében. Az adatok arra utalnak, hogy a hangképzés sajátosságai és a felismerés között szoros az összefüggés. Összességében kimondható, hogy az észlelést az eredeti magánhangzó minősége - vagyis az adott artikulációs mozgássor következtében létrejövő akusztikai formáció - meghatározza, ezen belül legnagyobb mértékben az első formánsok értékének, majd az F2-knek, végül az időtartamnak van szerepe.

5. táblázat: Az akusztikai adatok és a percepció összefüggései

Magán- hangzó	Korreláció	r	Értelmezés
o	idő/percep.	0,17891	nincs korr.
	F1/percep.	0,31886	közepes korr.
	F2/percep.	-0,74000	erős korr.
	idő/F1	0,22106	nincs korr.
	idő/F2	-0,20817	nincs korr.
ɔ	idő/percep.	-0,38644	közepes korr.
	F1/percep.	-0,52427	erős korr.
	F2/percep.	0,40336	közepes korr.
	idő/F1	0,37416	közepes korr.
	idő/F2	-0,2376	nincs korr.
a:	idő/percep.	0,43666	közepes korr.
	F1/percep.	0,76717	erős korr.
	F2/percep.	-0,63991	erős korr.
	idő/F1	0,43601	közepes korr.
	idő/F2	-0,40085	közepes korr.
e	idő/percep.	-0,42007	közepes korr.
	F1/percep.	-0,50174	erős korr.
	F2/percep.	-0,42817	közepes korr.
	idő/F1	0,315221	közepes korr.
	idő/F2	0,294632	nincs korr.

Noha a hangsúlyos magánhangzók előfordulása értelemszerűen alacsony; megnéztük, hogy a kapott adatok mutatnak-e jelentős eltérést a hangsúly megjelenésének függvényében. A teljes anyagot 100%-nak véve az összes hangsúlyos magánhangzó aránya 24,2%; ezt tekintve 100%-nak 30,9% az [o], 15,6% az [ɔ], 26% az [a:] és 27,5% az [e] esetében. Változnak az arányok, ha megnézzük, hogy a különféleképpen észlelt magánhangzóknál milyen arányú a hangsúlyosok megjelenése (6. táblázat).

6. táblázat: A hangsúlyos és a hangsúlytalan magánhangzók azonosítása

Magánhangzó (eredeti/észlelt)	Előfordulás (%)	
	hangsúlyos	hangsúlytalan
o/o	28,5	71,5
o/ə	33,3	66,6
ɔ/ɔ	16,6	83,4
ɔ/ə	10,9	89,1
a:/a:	50	50
a:/ə	-	100
e/e	42,8	57,2
e/ə	12,2	87,8

A hangsúlyosokat - az arányukon belül - nagyobb mértékben ismerik fel a kísérleti személyek az eredeti magánhangzónak, különösen jellemző ez az [a:] és az [e] magánhangzókra. Nincs szignifikáns különbség az [a] és az [o] esetében (utóbbinál a svaként azonosított hangok kissé nagyobb arányúak hangsúlyos helyzetben). Sajátosan alakulnak az időtartamértékek a hangsúlyos helyzet függvényében (7. táblázat).

A feltételezésünk az volt, hogy a hangsúlyos magánhangzók fizikai időtartama lényegesen meghaladja a hangsúlytalanokét, noha az [ɔ] és az [e] időtartamának alakulása abszolút szóvégi helyzetben befolyásolhatja ezt (vö. Kassai 1982). A veláris középső és alsó nyelvéllátású hangok esetében a feltételezés nem teljesen igazolódott: az eredeti

hangminőség megőrzésekor az időtartam nagyobb volt hangsúlytalan helyzetben, mint hangsúlyosban; ugyanakkor a sva azonosításnál jelentősen csökkent a hangsúlytalanok időtartama. Az [a:] esetében a hangsúlytalan helyzetűeké a rövidebb (megjegyzendő, hogy itt korlátozottak az adatok). Az [e] esetében alig jelentkezik különbség az időtartamban az eredeti magánhangzó azonosításakor, de a sva azonosítás esetén sem szignifikáns az időtartambeli különbség.

7. táblázat: A fizikai időtartamok hangsúlyos/hangsúlytalan magánhangzóknál

Magánhangzó (eredeti/észlelt)	Fizikai időtartam (ms)	
	hangsúlyos	hangsúlytalan
o/o	56,0	71,6
o/ə	72,5	61,0
a/a	85,5	94,2
a/ə	87,8	66,6
a:/a:	122,0	87,6
a:/ə	-	87,5
e/e	104,2	101,2
e/ə	71,2	66,3

A másik elemzési lehetőség azt mutatja meg, hogy a hangsúlytalan helyzetű hangok időtartamérték-átlagai - az [a:] kivételével - kisebbek a semleges magánhangzó azonosítása esetén. A hangsúlyos helyzetűeknél nem volt különbség az [ɔ]-nál és fordítottnál jelentkezett az arány az [o]-nál. A formánsok értékei nem mutattak szignifikáns eltérést a hangsúlyos helyzettől függően anyagunkban egyetlen magánhangzó esetében sem (ez adódhat a hangsúlyos hangok relative kis elemszámából is).

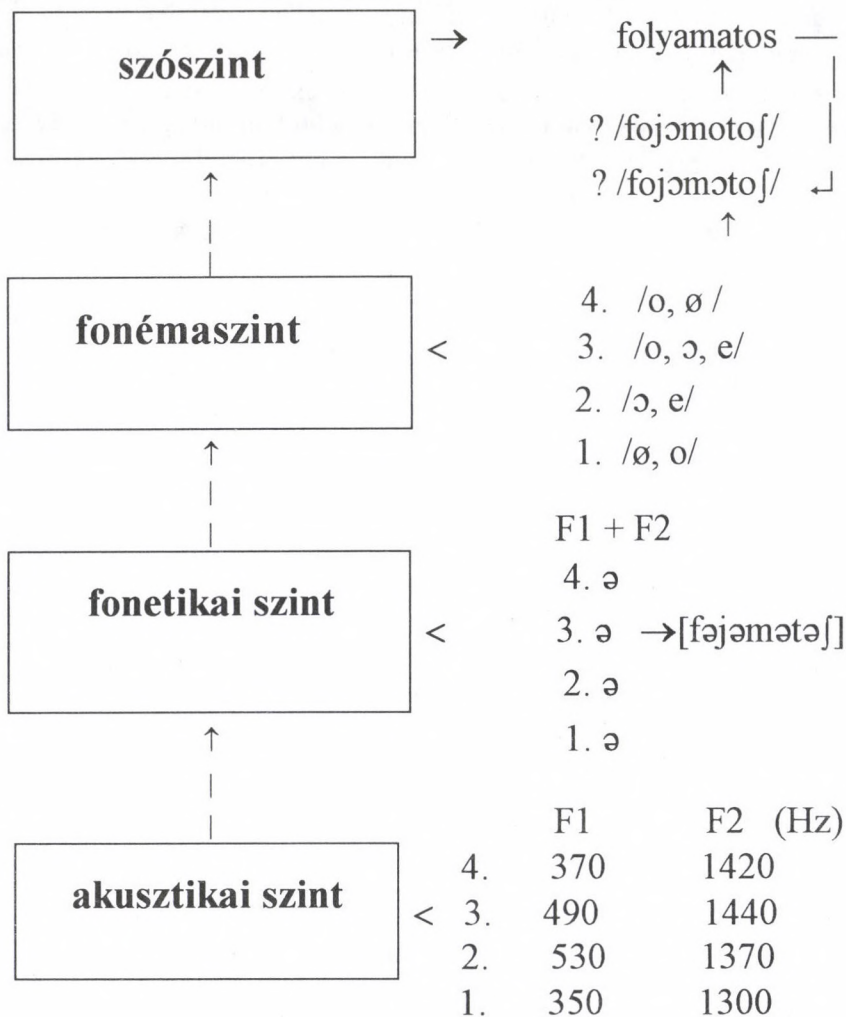
Megvizsgáltuk, hogy az adatközlő által ejtett [ø,ø:] magánhangzók az F1/F2 területet tekintve milyen viszonyban vannak az ugyanígy ábrázolt magánhangzókkal sva-azonosításkor. Az eredmények azt mutatják, hogy a svanak azonosított magánhangzók F1/F2 szerkezete többkevesebb mértékben fedi az eredeti palatális középső nyelvállású magánhangzók akusztikai tartományát, de jellegzetes módon egyben

megőrzik az eredeti magánhangzóra jellemző struktúrát is. Mind az első, mind a második formáns értékei magukban foglalják az eredeti hangminőségre jellemző értéket. Például amíg az eredeti [ɔ] helyén jelentkező sva F1-e 500 Hz táján jelentkezik (korrekt azonosításkor a hang első formánsának határértékei: 483-600 Hz), addig az [a:] helyén jelentkező sva átlagértéke 600 Hz felé tolódik (korrekt azonosításkor a határértékek: 600-760 Hz). Az F2 értéke az [e] helyén jelentkező sva esetén 1800 Hz-ig míg az [ɔ] helyén lévő semleges magánhangzók esetében csak mintegy 1600 Hz-ig tolódik ki.

Az 5. ábra - mindezek ismeretében - a szófelismerés egy lehetséges modelljét ábrázolja az elemzett spontán beszéd esetében. A *folyamatos* szó feldolgozásának folyamata látható a lexikális hozzáférés mechanizmusában a beszédészlelés egyes szintjein át a szóazonosításig. Az akusztikai szinten - az egyszerűség kedvéért itt csupán - az első két formáns frekvenciaértékeit tüntettük fel. A fonetikai szint az akusztikai elemzési eredmények alapján azonosítja a beszédhangokat, valamennyi magánhangzó esetében a semleges svat. A fonémaszint az összes eddigi eredmény alapján a lehetséges "eredeti" magánhangzó-fonémákat sorolja fel; majd a magyar fonotaktikai szabályok alkalmazása révén kiválasztja a kizárólag lehetségeseket. A szószint döntési folyamatai a mentális lexikon egységei alapján pedig végérvényesen meghatározzák az adott szót.

Következtetések

Az adatközlőnk spontán beszédanyagának szavaiban lévő elemzett magánhangzók kétséget kizáróan igazolták a hipotézist: a beszélő/hallgató által teljes értékű magánhangzóként észlelt beszédhangok jelentős része valójában nem azonos a vélt magánhangzóval. Azt feltételeztük, hogy a semleges hangminőség irányába történő változás egyik oka a beszédtempó gyorsulása. A műszeres elemzés ezt csak részben támasztotta alá, mivel adatközlőnk beszédtempója a lassú és az átlagos artikulációs tempó között váltakozott, a hangminőségbeli átcsapások mégis jelentős mértékben realizálódtak. A beszédtempó önmagában tehát csak kevésbé okozhatta a változást. A másik feltételezésünk az okra vonatkozóan magának, a spontán beszédnek a ténye.



5. ábra

Úgy tűnik, hogy a kapott adatmennyiség, a semleges magánhangzók közel 50%-os megjelenése a spontán beszédben arra utal, hogy a folyamatos artikulációban, illetőleg a beszédtervezés spontaneitásában kell a valódi ok(ok)at keresünk.

Spontán beszéd esetén a gondolati tartalmak diszkrét nyelvi egységekbe (szavakba, szószerkezetekbe stb.) alakítása a meghatározó a beszédtervezési folyamatban. Ennek az a következménye, hogy az artikulációs mozgássorok megtervezésére lényegesen kevesebb idő és figyelem jut, az itt lezajló transzformációk gyorsabbak, mint például hangos olvasáskor (ahol valamennyi transzformáció minden szempontból lényegesen kiegyensúlyozottabb). Az előre nem tervezett kommunikáció esetén a gondolati tartalom diszkrét nyelvi egységekké alakítása áll a középpontban, ezért az artikulációs mozgássorokban történő kifejezés nagyon automatikus műveletsorozatként realizálódik. Más szavakkal megfogalmazva, a beszélő elsődleges célja a gondolati tartalom nyelvi megformálása, s csak ezután, tehát másodlagosan jelentkezik az aktuális meghangosítás feladata. Az elsődleges transzformációsorozat mindenkor egyedi, hiszen a legtöbb esetben korábban ugyanúgy még nem megfogalmazott nyelvi szerkezetek generálása történik. A generált nyelvi szerkezetek meghangosítása azonban - az artikulációs műveletsorozatnak legalábbis a szegmentális szintjén - ismételt, begyakorolt gesztusok láncolata. Példával illusztrálva az elemzett anyag alapján ez a következőt jelenti. Részlet az adatközlő hanganyagából, két hosszabb "gondolkodási", azaz beszédtervezési szünet között elhangzott szöveg (a központosítást csak a könnyebb olvashatóság érdekében alkalmaztuk):

"...vannak köztük kaland, vannak akciók, általában különböző típusú gyerekeknek készült, vannak olyanok, amik inkább feszültséglevezető folyamatos izgalmat biztosítanak..."

A szövegrészlet szintaktikai megvalósítása, az adott tartalomnak a jelen szerkezet(ek)kel történt létrehozása egyedi, egy adott helyzet adott körülményei között fogalmazódott meg. A szerkezet lexikai egységei, szintagmái azonban az artikuláció szintjén nagy valószínűséggel nagyon sokszor létrehozott műveletsorozatok. Nem feltételezhető, hogy az adatközlő ezt a mondatot valaha is elmondta már (de ugyanebben a nyelvi formában legalábbis nem sokszor), míg - másfelől - nem felté-

telezhető, hogy ne ejtette volna ki (sokszor) azokat a közlést alkotó szavakat, mint *vannak, készült, olyanok, izgalmat* vagy *gyerekeknek*. A gondolat nyelvi transzformálása tehát nem klisék sorozata, a kiejtés azonban már igen, még akkor is, ha az illesztés (az egymás mellé került szavak koartikulációja) ritkább előfordulású, esetleg első ízben történt meg így.

1981-es adatok szerint a német nyelvben a semleges magánhangzók 23,88%-ban, az angolban 22,99%-ban, a franciában pedig 7,6%-ban jelentkezik (Onishi). Ezek az értékek a nyelv hangzó változatának szabályos (fonémának megfelelő) sva-realizálásait tükrözik. Nincs adat arról, hogy ugyanezekben a nyelvekben spontán beszéd esetén miként változik meg a többi magánhangzó hangminősége, milyen arányban csapnak át azok a semleges magánhangzóba. A lexikális hozzáférés folyamatának a spontán beszédben azonban mindenképpen lényegesen különböznie kell az izolált szavak azonosításának vagy az olvasott szöveg azonosításának folyamatától.

Kísérleti eredményeink alapján jogosnak látszik a magyar magánhangzók minőségéről - ma már klasszikusnak tekinthető - állítás revidálása: a spontán beszédben a magánhangzók egy részének hangminősége megváltozik, a semleges sva akusztikai sajátosságait tartalmazza, s ez egyértelműen a spontán beszédkivitelezés klisészerű lehívási folyamatára utal. A jelenség nemcsak a hangsúlytalan, hanem a hangsúlyos magánhangzók esetében is jelentkezik, és csak részben függ az aktuális (fizikai) időtartamtól. Nyelvhelyességi és kiejtés-stilisztikai szempontból feltehetően nem-kívánatos változás eredményeit regisztráltuk.

Irodalom

Bolla Kálmán: Magyar hangalbum. In: Fejezetek a leíró magyar hangtamból. Szerk.: Bolla Kálmán. Akadémiai Kiadó. Budapest 1982, 165-173.

Fry, D. (szerk.): Acoustic Phonetics. Cambridge University Press. Cambridge 1976.

Gósy Mária: Beszédeszélelés. MTA Nyelvtudományi Intézete. Budapest, 1989.

Jones, D.: The Pronunciation of English. Cambridge University Press. Cambridge 1966.

Kassai Ilona: Sva-jelenségek a magyar beszédben. Magyar Fonetikai Füzetek 1. 1978, 92-96.

Kassai Ilona: Időtartam és kvantitás a magyarban. In: Fejezetek a leíró magyar hangtanból. Szerk.: Bolla Kálmán. Akadémiai Kiadó. Budapest 1982, 115-145.

Koopmans-van Beinum, F. J.: What's in a schwa? Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica* 51. 1994, 68-80.

Lazicius Gyula: Fonetika. Tankönyvkiadó. Budapest 1963.

Lindner, G.: Grundlagen und Anwendung der Phonetik. Akademie Verlag. Berlin 1981.

Magdics Klára: A magyar beszédhangok akusztikai szerkezete. *NyudÉrt.* Budapest 1965.

Molnár József: A magyar beszédhangok atlasza. Tankönyvkiadó. Budapest 1970.

Olaszy Gábor: A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. *NyudÉrt.* 121. Budapest 1985.

Onishi, M. (szerk.): A Grand Dictionary of Phonetics. The Phonic Society of Japan. Tokio 1981.

Szalai Enikő: Az [u:], [a:] és [i] hangok koartikulációs mezőiről. In: *Beszédkutatás* 95. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1995, 83-93.

Tarnóczy Tamás: A magyar magánhangzók akusztikai szerkezete. *Kir. Magy. Pázmány Péter Tudományegyetem.* Budapest 1941.

Tarnóczy, T.: Acoustic analysis of Hungarian vowels. *QPSR*, Január 1965. Speech Transmission Laboratory, KTH. Stockholm 1965, 8-12.

A kutatás a T 013863 sz. OTKA keretében készült.

A BESZÉLŐ SZEMÉLY AZONOSSÁGÁNAK MEGÁLLAPÍTÁSA SZŰK FREKVENCIÁS SZAVAK ALAPJÁN

Nikléczy Péter
MTA Nyelvtudományi Intézet

Bevezetés

Az elmúlt néhány évben a személy azonosságának hang alapján történő megállapítására egyre nagyobb társadalmi igény mutatkozik: a) a kriminalisztikában - a beszédjellemzők felhasználhatók egyes eljárások során kiegészítő azonosítási vizsgálatként, illetve b) biztonságtechnikai rendszerekben - az egyén személyazonosságának megállapítása szavak alapján lehetőséget nyújt olyan rendszerek kialakítására, ahol a beszélő által ejtett hangsor mint "ujjlenyomat" szerepel. Ezek olyan kombinált azonosítók, amelyek az elhangzó szót, mondatot az előre felvett beszédanyag alapján hasonlítják össze a beszélővel. Mindezek alapján a modern beszélőazonosításnak két iránya alakult ki, amelyekben közös, hogy az alapot a beszéd egyéni akusztikai jellemzői adják (Hollien 1990). Az első a voltaképpeni személyazonosítás, amikor n számú lehetséges beszélő közül kell kiválasztani a rendelkezésre álló rögzített beszédanyag alapján annak eredeti közlőjét. A másodikat a szakirodalom beszélő-igazolásnak (speaker verification) nevezi (Schroeder 1985, 94); ez voltaképpen egy bináris döntés, amikor arra kell választani, hogy az eltárolt beszéd közlője azonos-e a bejelentkezővel.

Elengedhetetlen tehát, hogy minél rövidebb időn belül rendelkezünk egy olyan biztosnak mondható eljárással, ahol a keresendő személy kilétét rövid időtartamú és technikailag rossz minőségű hangfelvétel alapján is megfelelő biztonsággal meg tudjuk határozni (Nolan et al. 1995). Kísérleteinkkel e cél elérése érdekében a voltaképpeni személyazonosítás problémaköréhez csatlakozunk.

A beszélő személy kilétének meghatározása még technikailag jónak mondható hangfelvételek összehasonlításakor sem egyszerű feladat. Az első és legfontosabb kérdés az, hogy az ember vajon képes-e úgy beszélni, hogy az mindenki mástól különbözzön, illetve igaz-e az, hogy

az egyes beszélő személyek közötti változatosság mindig nagyobb, mint az egyén beszédében meglévő spontán változás. Ha az előbb feltett kérdésekre igenlő választ adunk, továbbra is problémák a következők: a beszéd akkor is alkalmas-e az azonosság megállapítására, ha 1. a rögzített beszéd időtartama relatíve rövid, 2. technikai színvonala rossz és 3. az összehasonlítandó hanganyagok különböző technikai körülmények (pl. stúdiófelvétel, ill. telefonos üzenetrögzítő) között lettek felvéve. Ezen kérdésekre nehéz egyértelmű választ adni, ugyanis a felsoroltakon kívül még számos egyéb tényező befolyásol(hat)ja az azonosító vizsgálat elvégzését. Az eredmény függ az elhangzó szövegtől (a hanganyag tartalmaz-e azonosításra alkalmas szóanyagot), a beszélő személy esetleges szándékos hangelváltoztatásától, a hangfelvételi körülményektől, a lehetséges beszélők számától (hány személyből kell kiválasztani a kérdéses személyt). Általánosan kimondható, hogy akkor a legnehezebb a beszélő személy azonosítása, ha sok a lehetséges személyek száma és korlátozott elemezhető beszédminta áll rendelkezésre. A személyazonosítás társadalmi igénye többnyire a telefonon érkező közlések elemzését kívánja meg, ezért szükségesnek láttuk, hogy további vizsgálódásainkat a szűk frekvenciatartományú beszédminták analizálására összpontosítsuk.

A kísérlet leírása

Kísérletsorozatot végeztünk annak megállapítására, hogy a szűk frekvenciás beszéd (200-2900 Hz) mely része tartalmazza a beszélő személy azonosításához leginkább szükséges akusztikai komponenseket. Az eddigi gyakorlat és a tapasztalati vizsgálatok azt mutatják, hogy az ilyen frekvenciatartománnyal rendelkező hangfelvételeknél a beszélő személyt a hallgató csak akkor ismeri fel, ha az közvetlen ismerőse, vagy ha az illető személlyel már többször beszélt telefonon. Különösen nehéz az azonosítás, ha az illető személy nő vagy gyerek, mivel az ő beszédük frekvenciatartománya a magasabb Hz-értékek felé tolódik el, így a telefon által áteresztett akusztikai tartomány nagyobb károsodást szenved a beszédfrekvenciák területén.

A kísérletben tizennyolc egyetemista lány vett részt, életkoruk 20 és 23 év között volt. Közülük tíz budapesti, nyolc pedig vidéki városból

származik; nyelvjárást nem beszéltek. Egyikőjüknek az általunk megadott telefonszámot fel kellett hívnia, és a zárt borítékban előre megkapott, gépelt szöveget a telefonba beolvasnia. A 18 lány maga döntötte el, hogy ki lesz a telefonba beszélő személy; kérésünkre ezt velünk nem közölték; a megkapott szöveget pedig csak a telefonálásra kiválasztott személy ismerte. A szöveg a következő volt: *"Jó napot kívánok! Azt szeretném bejelenteni, hogy ma este az Operettszínházban az előadás elmarad"*. Ezt az anyagot a Fonetikai Laboratórium saját vonaláról rögzítettük PMX88A típusú üzenetrögzítő segítségével. Ezután a kísérletben részt vevő összes (18) személy a Fonetikai Laboratóriumban az adott szöveget stúdiókörülmények között felolvasta, a szövegeket Studer A80-as stúdiómagnetofonra rögzítettük. A felolvasást minden esetben háromszor megismételtettük.

A cél az volt, hogy a két hangfelvétel összehasonlítása alapján a 18 személy közül meghatározzuk azt a személyt, aki a kérdéses telefonbejelentést megtette. A kísérlet annak megállapítását tűzi ki célul, hogy **a formánsstruktúra alapján történő azonosításhoz elégséges paraméterekkel nem rendelkező hangfelvételeknél a beszélő személyek azonosítása műszerek segítségével egyáltalán lehetséges-e, illetve hogy az egyén korlátozott felhangstruktúrája felhasználható-e az azonosításra.**

Módszer

A kísérlet kiindulópontját az a már említett tény adja, hogy a jól ismert vagy többször hallott személyt a hallgató bizonyos idő elteltével (telefonvonalon keresztül is) azonosítani tudja. Ez azt a feltevést erősíti meg, hogy a normál beszéd egy szűkebb frekvenciatartománya is tartalmaz olyan információt, ami a beszélő személy azonosításához elegendő. A kérdés az, hogy melyek ezek az akusztikai információk, valamint hogy ezek az információk jelfeldolgozó analizáló műszerekkel egyértelműen kimutathatók-e?

Az elemzéshez KAY gyártmányú 4300B típusú digitális jelfeldolgozót használtunk. Ez a készülék lehetővé teszi a hanganyag többféle analizálását, digitális szűrését, illetve a hangösszetevők numerikus megjelenítését.

A hanganyagok digitalizálásához 12000 Hz-es mintavételezési frekvenciát választottunk. Mivel ez a sebesség 6000 Hz-es elemzési tartományt enged meg, és viszonylag nem foglal le nagy tárhelykapacitást, jelen esetben ideálisnak mondható. A telefonvonalon keresztül rögzített hangfelvétel frekvenciatartománya 200 és 2900 Hz között van. A rögzített hangfelvétel zajjal erősen fedett. A 2900 Hz-es tartományban az 1200 és 2000 Hz között megjelenő intenzív zaj miatt az elemzés az 1200 Hz alatti frekvenciatartomány "információértékének" vizsgálatára irányult. Mivel a felhangszerkezet képbeli megjelenítésére és vizuális összehasonlíthatóságára a spektrografikus elemzés a legalkalmasabb, így ezt használtuk, és az általunk beállított alábbi értékekkel végeztük:

Frame Length: 23 Hz (keskeny sávú)

Display Range: 0-1200 Hz

Log or Linear Display: linear

A többi érték az alapbeállítást mutatja.

A 23 Hz-es keskeny sávú szűrőállítás beállítása lehetővé teszi a felhangok kirajzoltatását és az intonációs szerkezet megjelenítését. Ha a hangok akusztikai paramétereit nézzük, akkor elvileg a magánhangzók első, második és harmadik formánsának frekvenciaértékei többé-kevésbé beleesnek ebbe a sávba. A mássalhangzók közül a nazális hangok jól, a zöngés zárhangok csak korlátozottan elemezhetők. A zörejangok azonban a 2900 Hz-es vágás miatt nem nyújtanak értékelhető információt. Az átviteli sáv korlátozott értékén kívül még további két tényező nehezíti meg a beszélő személy azonosítását. Az egyik az, hogy a rögzített hangfelvételek zöme nagyon rossz jel/zaj viszonytal rendelkezik. A másik pedig az, hogy az átvitt frekvenciasávon belül a telefonvonal erősítőrendszerének karakterisztikája elsősorban a második és bizonyos mértékig a harmadik formánsok intenzitását módosítja.

Kutatási előzmények

A többéves tapasztalat azt mutatja, hogy a formánsok elhelyezkedésének összehasonlítása szűk frekvenciasávban csak abban az esetben ad értékelhető információt, ha az elmondott szöveg megfelelő időtartamú (legalább 60 s), és egyes hangok, hangkapcsolatok mindkét hanganyagban többször előfordulnak. Ez esetben a hangok tiszta fázisainak össze-

hasonlítása, valamint a hangkapcsolatok elemzése mutathat egyénre jellemző összetevőket.

A korábban elvégzett vizsgálatok alapján azt állapítottuk meg, hogy - az elmondottakat figyelembe véve - a felhangstruktúra a 200 és 1200 Hz-es tartományban analizálható. Ez az a rész, ahol az átviteli berendezések nagyjából lineárisnak tekinthetők. Ez a tartomány a magánhangzók közül az [ɔ, o, ɒ, u, u:] hangok első két, az [a:, i, i:, y, y:, e:, e, ø, ø:] hangok első formánsát tartalmazza. Mivel a formánsok tulajdonképpen a hangszalagok rezgése által gerjesztett felhangok szájüreg-rezonátorokban felerősített részei, így egy formáns több felhangot is magában foglalhat. Ezek a több felhangot magában foglaló (a rezonátorüreg által kiemelt felhangok) formánsok szerkezetükben azonban nem homogének. Ennek a különbözőségnek a kimutatása teheti lehetővé a beszélő személy azonosítását, illetve kizárhatóságát. Ahhoz, hogy egy formánsstruktúrán belüli felhangváltozást ki tudjunk mutatni, elsősorban az elemző módszert kell megfelelően kiválasztanunk. Arra kell törekednünk, hogy a mérendő értékek tisztán leolvashatók, illetve azok ismételt elemzés esetén könnyen reprodukálhatóak legyenek.

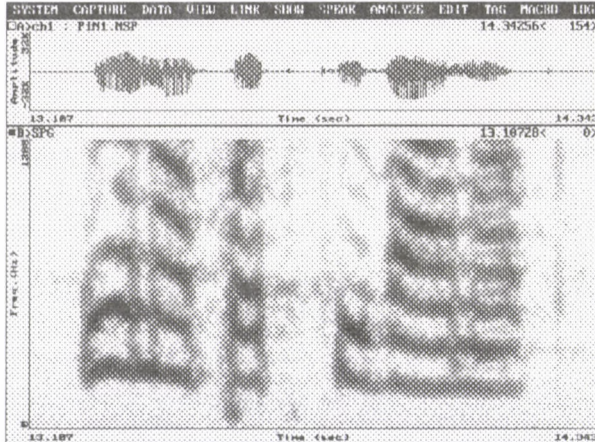
A kísérleti anyag elemzése

Az üzenetrögzítőre felvett azonosítandó hanganyagot (a.h.), valamint az egyes személyek hanganyagát (a személyre utaló betűjelekkel ellátva) digitalizáltuk és külön-külön fájlban rögzítettük. Ezáltal minden egyes hanganyag az a.h. felvétellel és egymással is összehasonlíthatóvá vált. Az azonosítandó hanganyag szubjektív, "füllel történő" többszöri meghallgatása után az első feladat azoknak a részeknek a kiválasztása volt, amelyek a beszélő személyek által feltételezhetően használt gyakori kifejezések. Ilyen volt a "*Jó napot kívánok*" hangsor. Ez a három szóból álló köszönés a mindennapi életben gyakran használt, tehát a beszélőnél motorikusan beidegződött, kimondásánál a szegmentális és szupraszegmentális tényezők valószínűsíthetően állandónak mondhatók.

Az első kísérletsorozat célja tehát annak megállapítása volt, hogy a kísérleti személyek e közlést a háromszori ismétléskor mennyire mondják azonosan. Az elvégzett személyenkénti 3, összesen 54 teljes spektrofizikus vizsgálat alapján (1620 mérés) megállapítottuk, hogy az e-

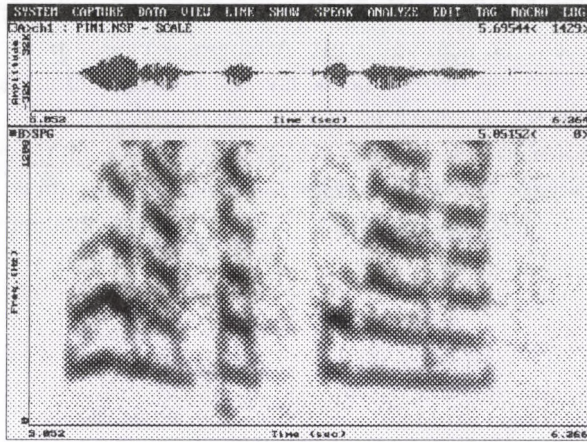
gyes személyek beszédének felhangszerkezete a többszöri bemondás esetén lényegesen nem változott. A hangok közötti átmenetek vonulatai nagyából azonosak maradtak, a hangokon belüli intonációs (mikrointonáció) változások iránya ugyancsak azonosnak mondható (vö. 1.a, 1.b ábra). Az 1.a és 1.b ábrán az KA jelű személy által néhány perces időkülönbséggel ejtett *"Jó napot kívánok"* hangsor spektrogramja látható. A hangsor oszcillografikus képe a felső ablakban, a spektrografikus képe az alsó ablakban látható. A hangszíneképeken a különböző hangok jól elkülöníthetők, esetleges hasonlóságuk vagy különbözőségük egyértelműen leolvasható. Mindkét hangsor kezdeti szakaszának dallamvonala emelkedő, majd ereszkedő jellegű. A felhangszerkezet frekvenciaértékei azonosak, az intenzitásmaximumok időben hasonló helyeken jelennek meg. Tehát a két hangsor az elemzési értékek alapján azonos személytől származtatható.

Az artikulációs mozgások akusztikai következményei tehát invariáns tendenciát mutattak.



1. a ábra

A KA jelű személy által ejtett *"Jó napot kívánok"* hangsor spektrogramja az első bemondáskor

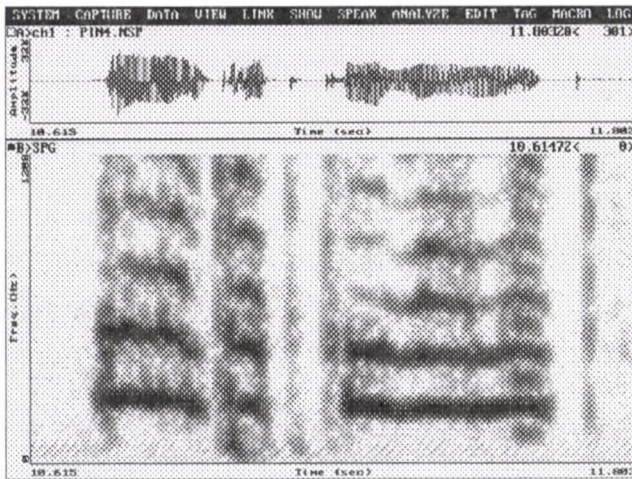


1.b ábra

Az KA jelű személy által ejtett "Jó napot kívánok!" hangsor
spektrogramja a második bemondáskor

Ezután annak eldöntése volt fontos, hogy ezen szűk frekvenciasávban a különböző beszélők által ejtett azonos hangsorok között van-e jól elkülöníthető különbség. Tekintettel arra, hogy a bejelentő személy esetében a hangsor csak egyszer hangzik el, fontos, hogy: a) ez a különbség jól kimutatható legyen, és b) az egyes beszélők által elmondott hangsorok egyike mutasson csak hasonlóságot az inkriminált hanganyaggal. Az elemzett adatok alapján kimondható, hogy a 17 x 14 (összesen 868) hang egymással összehasonlítva felhangstruktúrájában és mikrointonációjában valóban jelentős különbségeket mutat (2. ábra)

Minderre példa a 2. ábrán látható TI jelű személy bemondásáról készített spektrografikus kép. Ha a hangszíneképet összehasonlítjuk az 1. ábra a. vagy b. képével, szembetűnő a vizuálisan is érzékelhető különbség. Nemcsak az intonációs szerkezetben, hanem a hangok időtartamában is eltérés mutatkozik (vö. 1. táblázat). Az elemzések azt mutatják, hogy az egyes személyek szavai közötti különbség kimutatható. Abban az esetben tehát, ha az azonosítandó személy a lehetséges személyek között van, akkor az hasonló analízis módszert alkalmazva szűk frekvenciatartományban is kiválasztható.



2. ábra

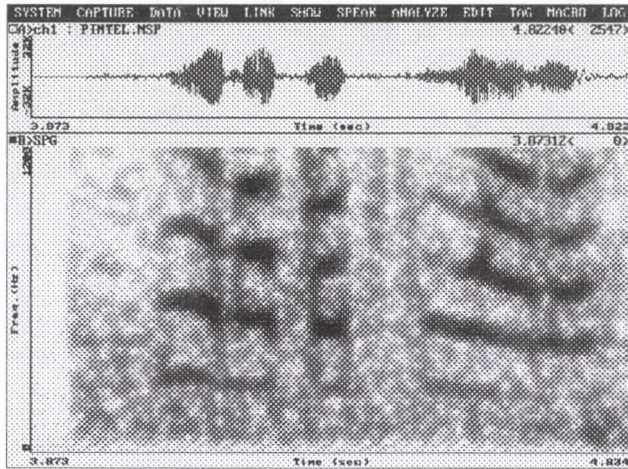
A TI jelű személy "Jó napot kívánok!" hangsorának spektrogramja

1. táblázat: A "Jó napot kívánok!" hangsor szótagjainak időtartamátlag 4 bemozdó esetében (adatok s-ban megadva)

hangsor	"Jó"	"na"	"pot"	"ki"	"vá"	"nok"
telefonáló	0,095	0.079	0.165	0.131	0.106	0.129
KA jelű bemozdó	0.1	0.123	0.222	0.153	0.172	0.228
PK jelű bemozdó	0.092	0.102	0.188	0.136	0.183	0.212
TI jelű bemozdó	0.106	0.092	0.189	0.180	0.125	0.236

A 3. ábrán a telefonvonalról rögzített hangfelvétel spektrografikus képe látható. A hangsor időtartama 0,705 s, meglehetősen rövid. A 200 és 1200 Hz közötti spektrografikus kép azonban a felhangstruktúrát értékelhetően mutatja. Jól látható a 250 Hz alatti rész gyenge intenzitása, valamint az, hogy az e fölötti rész megjelenése több intenzív felhangot

mutat a stúdiófelvételekről készített spektrogramokénál. (Ezt a korábban említett telefonvonalai erősítőrendszer okozza. E különbséget az összehasonlításnál majd figyelembe kell venni.)

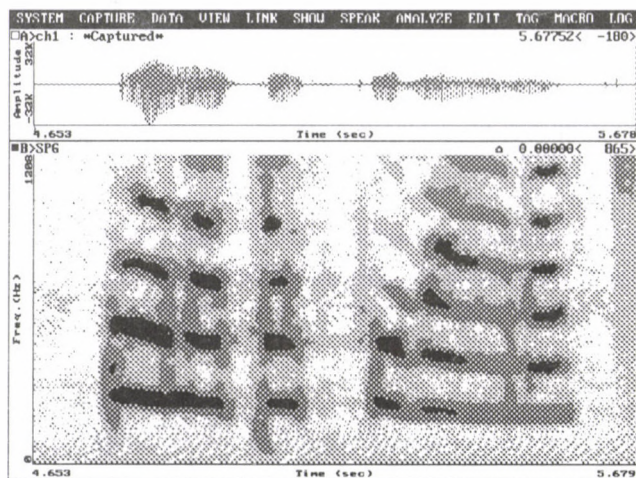


3. ábra

A telefonvonalról rögzített, azonosítandó "Jó napot kívánok" hangsor

A kísérleti személyek bemondásairól készített spektrografikus képek kiértékelése alapján (vö. KA és TI korábban már bemutatott hangszínképét), a numerikus és vizuális összehasonlítások eredményeképpen a **telefonos bejelentő személyt a PK jelű kísérleti személlyel** azonosítottuk. Az azonosítás pozitív eredményt hozott, mivel a kísérletben résztvevők később igazolták a bejelentő személy kilétét. A 4. ábrán látható PK jelű bemondó spektrogramja az, aminek hasonlósága alapján az illető személyt azonosnak találtuk a telefonos bejelentő személlyel. A két felvétel természetesen tökéletes azonosságot nem mutat. Nem is mutathat, hiszen a két hangfelvétel között időben majdnem három hét telt el. Ha a numerikus értékeket a jelfeldolgozó ASL (Analysis Syntesis Lab.) szoftverével analizáljuk, majd a kapott értékeket numerikus formában kiírjuk, akkor különbséget főleg az alaphang és ezáltal az első

formáns elhelyezkedésében találunk. Ezért a kísérlet egyik legfontosabb megállapítása az, hogy a beszélő személy azonosítását szűk frekvenciás szavak alapján elsősorban nem a számszerű értékek megegyezésére kell alapozni, hanem a teljes spektrogram képi azonosságát kell figyelembe venni. Ez azonban csak a már említett beidegződött, motorikusan ejtett kifejezések összevetésénél mondható ki. A beszélő személy azonosításában a spektrografikus eljárással tehát nem a numerikus azonosság, hanem az artikuláció jellegzetes akusztikai megjelenésének összes, vizuálisan azonosítható sajátossága a döntő.



4. ábra

A PK jelű személy bemondásában a "*Jó napot kívánok*" hangsor

Az ábrán jól látható, hogy az azonosításhoz már a szószintű összehasonlítás is elégséges. A kísérlet végeredményben azt mutatta meg, hogy bizonyos körülmények között a beszélő személy azonosságát szűk tartományban is ki lehet mutatni. Természetesen a módszer nem alkalmazható mindenkor és minden esetben, hiszen sokszor még ez a szűk frekvenciasáv sem áll torzítatlanul az összehasonlítást végző személy rendelkezésére.

Szubjektív tesztelési adatok

A műszeres vizsgálat befejeztével ellenőrző kísérletet végeztünk annak megállapítására, hogy az azonosság ilyen nagyszámú személy esetében füllel mennyire meghatározható. 10 kijelölt személy a telefonos bejelentést és a 18 kísérleti alany beszédét véletlenszerű sorrendben hallgatta meg oly módon, hogy a telefonos közlés minden harmadik lehetséges beszélő mintája után megismétlődött. A lehallgatók feladata az volt, hogy a 18 bemondót három csoportra különítsék el. Az első csoportba a kizárható személyeket, a második csoportba a szóba jöhető személyeket, a harmadik csoportba pedig a leginkább azonosnak vélt személyeket írják fel. A kísérletet kétszer ismételtük meg úgy, hogy a bemondó személyek sorrendjét megváltoztattuk. A kísérlet előre nem várt eredménnyel zárult. A hallgatók a műszeresen azonosított PK jelű személyt mindkét esetben a kizárható személyek közé sorolták. Véleményünk szerint ennek egyedüli lehetséges oka az, hogy a rendelkezésre álló rövid közlés nem elegendő a személyazonosításra akkor, ha a lehallgatók a kérdéses beszélőt korábban nem ismerték. Ekkor ugyanis nincs meg az agyban az a tárolt minta, amelynek alapján az összehasonlítás elvégezhető lenne.

Irodalom

- Schroeder, M. R.: Speech and Speaker Recognition. Karger. Basel 1985.
 Nolan, F.–Hollien, H.–Braun, A.–Broeders, A.P.: Forensic phonetics. In: Proc. of 13th ICPhS. Eds. K. Elenius–P. Branderud. Vol. 3. 1995, 130-162.
 Hollien, H.: The Acoustics of Crime. The New Science of Forensic Phonetics. Plenum Press. New York 1990.

Ez a kutatás a T013863 sz. OTKA keretében készült.

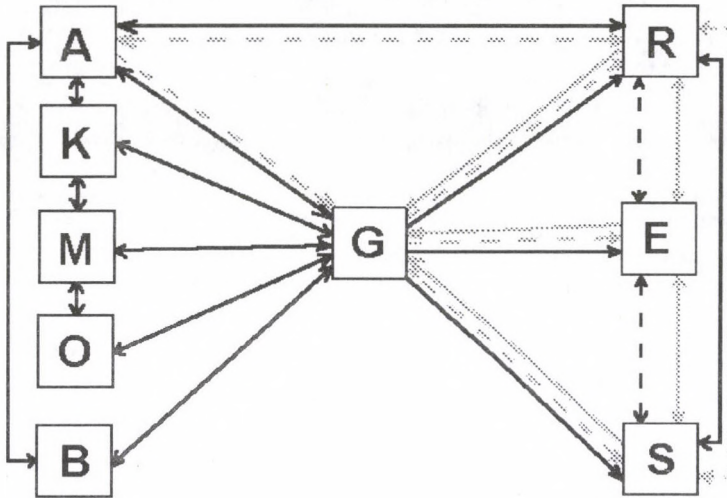
FONETIKAI JELENSÉGEK EGY KÉTÉVES GYERMEK NYELVÉBEN

Kovács Andrea
MTA Nyelvtudományi Intézete

Bevezetés

A hazai gyermeknyelvi irodalom túlnyomó többsége egy-egy gyermek nyelvi fejlődéséről kíván képet nyújtani. E tanulmány is kapcsolódik ehhez a hagyományhoz abból a szempontból, hogy egyetlen megfigyelt gyermek a vizsgálati személy, ellenben célja nem az, hogy általános nyelvi fejlődésrajzot adjon, hanem egy adott időszak hang-állományát és fonetikai jellegzetességeit kívánja bemutatni és rendszerezni. A vizsgálati módot a korábbi kutatások már igazolták (Ronjat 1913; Pavlovich 1920; Leopold 1949; Imedadze 1960; S. Meggyes 1971; Kassai 1981; Gósy 1984). Az egyes gyermekek egyéni nyelvi fejlődése között tapasztalt óriási eltérések ellenére kimutathatóak általános tendenciák.

A tanulmányban fiam, G. kétéves kora körüli nyelvállapotát kívánom analizálni. A részletes elemzés előtt ki kell térni G. nyelvi környezetére (1. ábra), ami a nyelvelsajátítás háttéréül szolgál. G. sajátos nyelvi környezetben, két nyelv állandó hatása alatt fejlődik. Helyzete azért mondható sajátosnak, mert környezetében mindenki beszél magyarul, lengyelül azonban csak négy személy (R., A., E., S.), s közülük rendszeres kommunikációra csupán hárman használják a lengyelt (R., E., S.), G.-vel pedig csak az apa áll mindennapos verbális kontaktusban lengyel nyelven. A szülő-gyermek kapcsolat kommunikációs stratégiája a Ronjat-séma szerinti "egy személy - egy nyelv" alapelv, vagyis R. lengyel irodalmi köznyelven, A. magyar irodalmi köznyelven beszél N. G.-hez. Magyar környezetben R. lengyelül, a környezet pedig magyarul érintkezik a gyermekkel. Ugyanez a stratégia érvényesül a G.-E. és a G.-S. interakciók során (vö. 1. ábra).



1. ábra
N.G. nyelvi környezete

Jelmagyarázat:

A. - anya; R. - apa

K. - anyai nagypapa (csak magyarul érintkezik N. G.-vel)

M. - anyai nagynéni (csak magyarul érintkezik N. G.-vel)

O. - anyai nagynéni (csak magyarul érintkezik N. G.-vel)

B. - bölcsődei közösség (N. G. 0;11 kora óta bölcsődébe jár, ahol csak magyarul érintkeznek vele)

E. - apai nagynéni (csak lengyelül beszél N. G.-hez)

S. - apai nagynéni (csak lengyelül beszél N. G.-hez)

Magyar verbális megnyilatkozás: fekete szín

Lengyel verbális megnyilatkozás: szürke szín

Gyakori (általános) alkalom: folytonos vonal

Ritka (kivételes) alkalom: szaggatott vonal

A nyíl mindig a beszélőtől a hallgató felé mutat.

Ezek alapján G. nyelvi környezetének jellemzőit a következőkben lehet összefoglalni:

Szülők: az egyik szülő kétnyelvű, a másik egynyelvű, de jó szinten ismeri és beszéli a másik nyelvet is.

Szociális milió: a környezeti nyelv azonos az egyik szülő anyanyelvével.

Kommunikációs stratégia: a kétnyelvű szülő a nem-domináns nyelven, az egynyelvű szülő pedig az anyanyelvén (domináns nyelv) érintkezik a gyermekkel.

A szülők egymás közti kommunikációs stratégiája: a környezeti (domináns) nyelven érintkeznek egymással.

N. G. nyelvi helyzetéről 2;0,16 korától 2;8,4 koráig rendszeres időközönként készített magnetofon-felvételek segítségével rögzítettem adatokat. (Az életkor megadásánál a gyermeknyelvi szakirodalomban szokásos jelölést használom.) G. a vizsgálat kezdetén a szómondatok szintjén, az önálló beszédprodukció kezdeti szakaszában volt. Artikulációja pontatlan, néha bizonytalan, megnyilatkozásai gyakran kevésbé érthetőek. Az időszak végére G. az önálló beszéd jóval magasabb fokán állt: magyarul grammatikailag hibátlan összetett mondatok generalizálására is képes volt, lengyel nyelvi beszédprodukciója azonban a megfigyelés utolsó fázisában is csak a holo-frázisok, a spontán vagy provokált ismétlések, illetve kisebb részben a telegrafikus közlések szintjén stagnált. Mindennek ellenére valószínűnek tűnik, hogy G. tudatában a két nyelv egységes gondolkodási bázist alkot, ezért az elemzés vezérfonalát nem a nyelvek különbözősége adja, hanem a fonetikai jelenségek vizsgálata.

A megfigyelt időszakra, illetve időszakban G. nyelvében már jó-részt kialakult a magyar és a lengyel fonémarendszer. A még hiányzó hangokat hanghelyettesítéssel pótolta. A hangrendszer viszonylagos kialakultsága ellenére is számos bizonytalanság, pontatlanság, az egyes hangok nagyszámú artikulációs variációja fordult elő. Ezek a jelenségek, valamint a hanghelyettesítések egyes típusai nagyjából megfelelnek a gyermekkori fiziológiás (Vértes 1956) és nem fiziológiás beszédhibák bizonyos fajtáinak (Montágh 1993).

1. Magánhangzók

1.1. Hangállomány

A magánhangzókat G. viszonylag tisztán ejtette, csupán néhány hang képzése okozott gondot a megfigyelés első felében. Így ebben a

periódusban hangállománya - elemszámát tekintve - nem egyezett meg a köznyelvvél: hiányoztak a szájüreg elülső részében képzett, ajak-kerekítéses, középső és felső nyelvállású, rövid és hosszú magánhangzók, az [ø, ø:, y, y:]. A lengyel hangkészletből csak a veláris [ɨ]* okozott néhányszor artikulációs problémát, ez azonban minden alkalommal átmeneti nehézségnek bizonyult, pl. [buti~buti] 'cipő'. Minden magánhangzóra szinte kivétel nélkül érvényes, hogy a nyelvállást tekintve meglehetősen széles képzési területen realizálódtak. Különösen jellemző volt ez az illabiális palatális hangokra.

1.2. Hangtani jelenségek

Állandó jelenség volt a magánhangzók időtartamában való ingadozás (vö. S. Meggyes 1971). A hangok időtartam-változásaiban szerepet játszott az adott szóforma intonációja. Nevezetesen a hangsúly, a hanglejtés, a dallamvezetés bizonytalanságát a (többnyire szóvégi) magánhangzó hosszúságának megváltoztatásával próbálta ellensúlyozni, illetve intonációs hibáit korrigálni. Kérés, óhaj kifejezésekor jellemző volt az utolsó magánhangzó erőteljes megnyújtása, pl.: [hintɔ:] 2;0,25 'hintázni szeretnék', [teje:] 2;0,18., 2;2,6 'kérek teát', [hus'i:] 2;1,1 'husit kérek vacsorára', [in:i:] 2;0,18 'Szomjas vagyok, kérek inni'. Ez a jelenség még a nyelvileg hosszú magánhangzókat is érintette. Ilyenkor a hosszú magánhangzó időtartama a szubjektív megítélés szerint a kétszeresére is megnőhetett, pl.: [zi:z] 2;1,5 'víz', [e:n] 2;2,12 'én akarom csinálni'.

A magánhangzó-rövidülés gyakran azt jelezte, hogy nagyon határozottan, "ellentmondást nem tűrve" jelentett ki valamint. Ilyenkor a rövid magánhangzót még rövidebben ejtette ki, pl.: [t'yt'] 2;0,16 'ülj le = csücs', [jo] 2;2,17 'jó', [enim] 2;3,24 'enyém'.

Ez a jelenség természetesen elsősorban a holofrázisokat jellemezte. A szerves mondatok általánossá válása után a jelenség gyakorisága nagymértékben csökkent, ami azt jelzi, hogy a nyelvi elemek és eszközök funkciója nagyon rugalmasan változik a gyermek nyelvi fejlődésével párhuzamosan. A gyermek már felismeri bizonyos nyelvi elemek, formák szükségszerűségét a beszédben, s ha azt ő még nem képes hibátlanul realizálni, egy másik nyelvi eszközzel helyettesíti

vagy ezen eszköz segítségével tökéletesíti a rendelkezésére álló nyelvi alakzatot (vö. Kassai 1979).

A magánhangzókkal kapcsolatos egyéb jelenségek a gyermeknyelvi irodalomból ismert módon jelentkeztek, ezért részletes elemzésükkel e tanulmány nem foglalkozik, csupán - a teljesség kedvéért - megemlíti és néhány példával illusztrálja azokat. Jellemző volt bizonyos magánhangzóknek a köznyelvinél zártabb ejtése. Ez a változás elsősorban a palatális illabiális hangokat érintette: [ni:mɔ] 2;0,16 'néma', [dajita] 2;2,2 'daleko=messze', [nehi:z] 2;0,18 'nehéz'. Az /e/ fonéma gyakran az alsó, illetve legalsó nyelvállású, nyíltabb képzésű magánhangzóban realizálódott (vö. Gósy 1984 is): [tæt:o:] 2;1,1 'kettő', [k'æj] 2;2,18 'kell', [hæjipot:ej] 2;3,5 'helikoptert', [ig'hæn] 2;4,2 'igen'. Előfordult, hogy veláris helyett palatális magánhangzó jelentkezett: [hændos'] 2;0,16 'hangos', [hæjo:] 2;2,7. 'halló'. E két szóban a magánhangzó nyíltabbá válása is megfigyelhető: [ɔ] > [æ]. Érdekes azonban, hogy - a hanghelyettesítés révén - azonos beszédhangokból felépülő lexéma, a "hajó" sohasem fordult elő ebben a variációban, vagyis a "hajó" [ɔ] hangja sohasem palatalizálódott. E jelenség mögött két ok feltételezhető. Egyrészt a "halló" képzését befolyásolhatta a szintén ismert "hellő" szó hangalakja. Másrészt elképzelhető, hogy G. így próbálta ösztönösen elkerülni a homonímák kialakulását. Néhányszor előfordult a lengyel hangrendszer felső nyelvállású veláris illabiális magánhangzójának, az [î]-nak megfelelő palatális párjával történt helyettesítése: [buti] 2;0,16 'buty =cipő', [toji] 2;4,24 'tory = sínek', [vadoni] 2;5,8 'wagony = vagonok', [du:ʒi] 2;4,24 'duzy = nagy'. Itt nem zárható ki a magyar nyelv hatása. A megfigyelés első felében állandó jelenségnek számított, de gyakorlatilag mindvégig jellemző maradt a palatális hangok velárisal való helyettesítése. Elsősorban a problémát okozó [y], [y:], [ø], [ø:] esetében fordult elő: [ɛ:do:] 2;0,17 'erdő', [jepujo:] 2;6,17 'repülő', [dojod] 2;6,18 'dörög', [hu:to:] 2;7,8 'hűtő'. A labiális - illabiális párok felcserélésére csak ritkán került sor: [pət'es'] 2;7,0 'pötytös', [hændos'] 2;0,16 'hangos'. A gyermeknyelvi szakirodalomban ismert [ɔ] > [a] hangváltozásra is volt példa: [japa:tot] 2;6,21 'lapátot', [mada:jot] 2;4,20 'madárot = madarat' (S.

Meggyes 1971; Gósy 1984). A tárgyalt jelenségek összefoglalásául ismertetem G. fonémaállományát és azt, hogy az egyes fonémák mely beszédhangokban valósulnak meg:

/ɔ/ > [ɔ], [a], [ɔ:], ([o])	/o/ > [o], [o:]
/a/ > [a]	/o:/ > [o:], [o]
/a:/ > [a:]	/ø/ > [ø], [o]
/e/ > [e], [æ]	/ø:/ > [ø:], [o:]
/e:/ > [e:], [i:], [ei], [ie]	/u/ > [u]
/i/ > [i]	/u:/ > [u:]
/i:/ > [i:]	/y/ > [y], [u]
/î/ > [î], [i]	/y:/ > [y:], [u:]

A magánhangzókkal kapcsolatos fonetikai jelenségek között meg kell még említeni az illeszkedés néhány hibáját, amely végigkísérte ugyan a megfigyelés egész időszakát, de rendkívül ritkán fordult elő. Mindössze két névszói helyhatározóragot érintett: *-nál*, *-nél*, *-hoz*, *hez* és *-höz*. A szavak hangrendje a tévesztés szempontjából közömbösnek tűnik: [*tatus*'*hez*] 2;3,25 '*tatushoz*' = *apuhoz*', [*bo:ts*'*ina:j*] 2;3,10 '*bölcsinél*', [*de:dihoz*] 2;5,14 '*dédihez*'.

2. Mássalhangzók

2.1. Hangállomány

A mássalhangzók hibátlan képzése a beszédelsajátítás folyamán általában sokkal több gondot okoz, mint a magánhangzóké (vö. S. Meggyes 1971; Gósy 1984). Ez G. esetében sem volt másként. A megfigyelt időszak adataiban az artikulációs megoldás alapján a mássalhangzók három csoportját lehetett elkülöníteni:

a) Képzési nehézséget nem okozó hangok, amelyek a hanghelyettesítésekben nem fordulnak elő: [b], [p], [f], [m], [n]*, [p], [t], [h]*, [j]*. (A csillaggal jelölt hangok helyettesítésére elvétve akad példa a mintanyagban, de szisztematikus helyettesítési rendszert nem alkotnak, előfordulásuk alkalmi artikulációs pontatlanság.)

b) Artikulációs pontatlansággal realizált hangok, amelyek a hanghelyettesítésekben gyakran előfordulnak: [dʒ], [ʃ], [s], [v], [z], [ʒ]. Az alveoláris centrális spiránsok képzésében az életkorra jellemző fizioló-

giai pöszeség jelentkezett, amely a [ɕ] artikulációjára is hatással volt: [ɕʰ].

c) Kizárólag helyettesítő hanggal realizált hangok: [j], [l], [nʰ], [r], [c]. Az adatgyűjtés megkezdésekor ide tartoztak még a [ts], [tʃ], [g], [k] hangok.

2.2. Hangtani jelenségek

A nehézséget okozó mássalhangzók helyettesítése nem mutatott jelentős eltérést az egynyelvű magyar gyermekek beszédprodukciónál megfigyelt jelenségektől (vö. Vértess 1963; Mikes 1966; S. Meggyes 1971; A. Molnár 1978; Gósy 1978, 1979). Tekintettel a bőséges szakirodalmi háttérre itt csupán N.G. fonémaállományát és a fonémák beszédhang-realizálódását ismertetem:

/b/ > [b]	/p/ > [p], egyszer: [b]
/d/ > [d]	/t/ > [t], ritkán: [d], [p]
/ɟ/ > [dʰ]	/c/ > [t]
/g/ > [g], [d], [gʰ]	/k/ > [k], [t], [kʰ]
/z/ > [sʰ], [z], ritkán: [d]	/s/ > [sʰ], [h], ritkán: [j]
/ʒ/ > [ʒ], [z]	/ɕ/ > [dʰ], [ɕʰ]
/j/ > [j]	/h/ > [h]
/x/ > -, [h]	/ts/ > [tʰ], [sʰ], [ts]
/tʃ/ > [tʰ], [sʰ], [tʃ], [tsʰ]	/tsʰ/ > [tʰ], [sʰ], [tsʰ]
/m/ > [m]	/n/ > [n], ritkán: [m]
/nʰ/ > [n], ritkán: [j]	/β/ > -, [j]
/r/ > [j], ritkán: [β], [h], [n]	/ɕʰ/ > [dʰ], [ɕʰ]
/ʃ/ > [j], [sʰ], ritkán: [n], [j]	
/l/ > [j], ritkán: [h], [v], [n], [m]	
/v/ > [v], [β] ritkán: [j], [h], [z], [m]	
/f/ > [f], ritkán: [j], [h], [t], [v]	

2.2.2. Mássalhangzó-kapcsolódások

A mássalhangzó-kapcsolatok a megfigyelt időszakban már sok esetben nem okoztak kiejtési problémát. Erre elég néhány példát idézni: [hɔndosʰ] 2;0,16 'hangos', [dɔjɔmb] 2;3,17 'galamb', [embej] 2;2,7 'ember'. Két mássalhangzó szóbelseji vagy szóvégi pozícióban történő találkozásánál nem jelentkezett a beszédhangok kiejtési nehézsége, azon-

ban a fonotaktikai struktúrában változás történt. Az artikulációs problémát jelentő mássalhangó-torlódások feloldását G. négyféleképpen valósította meg.

a) Az egyik mássalhangzó teljes elhagyása és a másik mássalhangzó egyidejű megnyúlása:

C1C2 > C1:

C1C2 > C2:

Ezekben az esetekben jellemző a képzési nehézséggel járó hang elhagyása, valamint az a szabályszerűség, hogy a megmaradt mássalhangzó megnyúlik. Példák: [mad:a] 2;1,19 'Magda', [fyt:e] 2;3,29 'füstje', [ɔb:ɔt] 2;7,23 'ablak', [d'ejet:] 2;4,24 'gdzie jest = hol van', [dob:e] 2;6,26 'dobrze = jól van', [mekfot:ɔm] 2;8,0 'megfogtam', [n'it:i] 2;7,4 'nyisd ki', [mon:em] 2;2,1 'most nem', [mɔn:o:] 2;0,16 'magnó', [fyt:øj] 2;8,0 'füstöl', [pit:ɔ] 2;5,1 'Pista', [hit:is'] 2;0,16 'hisztis', [hot:ɔ] 2;1,29 'hozta'.

b) Az első mássalhangzó teljes elhagyása és a megelőző magánhangzó egyidejű megnyúlása:

VC1C2 > V:C2.

A mássalhangzó-torlódás ilyen feloldása két hang esetében volt jellemző. Az egyik az [r] + mássalhangzó, a másik az [l] + mássalhangzó hangkapcsolat. Ezekben az esetekben minden alkalommal a kritikus képzésű hangok, a likvida és a tremuláns estek ki az artikuláció során, s minden esetben a V1 pozícióból. A V1 pozíció előtt álló magánhangzó pedig megnyúlt, vagy nyelviileg hosszú volt. Ez a jelenség néhány magyar nyelvjárásban is ismeretes. Példák: [to:to:dmeg] 2;7,4 'kóstold meg', [mekpo:ba:jom] 2;8,4 'megpróbálok', [to:na:zok] 2;8,3 'tornázok', [s'ø:pit] 2;6,5 'szörpit', [tamvaj] 2;5,17 'tramvaj = villamos', [minda:t] 2;2,7 'mindjárt', [tø:te:t] 2;5,4 'körtét', [ɔ:tsɔ] 2;7,3 'arca', [bejefe:t] 2;7,3 'belefért', [βo:t] 2;2,2 'volt', [bø:ts'ibe] 2;3,28 'bölcsibe', [bo:dod] 2;2,7 'boldog', [s'ikejy:t] 2;3,30 'sikerült', [ts'ina:s'] 2;4,8 'csinálsz', [med'o:du:t.] 2;7,23 'meggyógyult', [do:dozot:] 2;7,23 'dolgozott', [uto:s'o:t] 2;7,23 'utolsót', [kija:ndu:ni] 2;6,13 'kirándulni', [kos'o:tɔm øs'e] 2;6,19 'koszoltam össze', [jeø:tøzøt:] 2;6,25 'leöltöztött', [bes'a:ni] 2;7,1 'beszállni', [føja:t] 2;7,3

'főállt', [e:bu:tɔm] 2;7,3 'elbújtam', [ejhy:t] 2;7,8 'elhűlt', [to:ni] 2;7,11 'tolni', [bepɔt:o:ni] 2;8,0 'bekapcsolni', [n'u:hɔts] 2;8,0 'nyúlhatsz', [pɔt:o:d] 2;8,2 'kapcsold', [mekɔgo:ni] 2;8,4 'megszagolni', [fogɔ:mɔm] 2;8,4 'fogalmam'.

c) A mássalhangzó-torlódás "összeolvadásos" feloldása:

C1C2 > C3.

Ebben az esetben egy harmadik mássalhangzó helyettesíti az eredeti két mássalhangzót. Ez a jelenség erősen emlékeztet a "felnőtt" köznyelv hangtani törvényére, az összeolvadásra. A két (vagy több) - képzési nehézséget okozó - hang kapcsolódásakor (pl.: -rg-, -gr-, -rk-, -kr-, -tr-, -fr-) az egyik hang kiesik, a másik pedig helyettesítő hang formájában realizálódik. Példák: [pɔjɔnt] 2;3,16 'Trabant', [te:toto:] 2;0,17 'krémtúró', [ɔt:itɔ] 2;1,22 'Afrika', [tid:is'] 2;1,22 'tigris', [fɔ:tɔs'] 2;0,16 'farkas', [tɔ:tɔ] 2;1,15 'tarka', [fɔ:tɔ] 2;1,15 'farka', [me:des'] 2;2,7 'mérges'.

Feltűnő, hogy e hangkapcsolatok egyik eleme minden esetben az [r]. Mivel ez az egyik legnehezebb képzésű hang a magyarban, természetes, hogy a mássalhangzó-torlódás feloldásakor ez esik ki. Az adatok az anyaggyűjtés kezdeti fázisából valók, amikor a [k] és a [g] (e hangkapcsolatok gyakori elemei) artikulációs problémát jelentettek (még VCV helyzetben is), ezért ezeket a hangokat G. a szokásos helyettesítő hangokkal realizálta. Ezt a jelenséget egyéb hangváltozás is kíséri. Nevezetesen vagy a mássalhangzó, vagy a megelőző magánhangzó megnyúlása. Éppen ezért ez a harmadik, c) típus csak bizonyos korlátokkal fogadható el önálló kategóriának, hiszen magában hordozza mind az a), mind a b) típus jellemzőit.

A c) megkülönböztető kritériuma a következő: egyszerre kell megvalósulnia a mássalhangzó-elhagyásnak és a hanghelyettesítésnek. Ennek pedig szükségszerű kísérőjelensége - természetesen a fonetikai pozíciótól függően - valamelyik hang időtartam-módosulása.

d) Kötőhang-betoldás, illetve nyílt szótag ejtésére törekvés:

C1C2 > C1VC2

A mássalhangzó-torlódás feloldásának ez a módja a megfigyelt időszak második periódusában volt jellemző. Elsősorban toldalékolt szóala-

kokban fordult elő. A kötőhang betoldásával párhuzamosan a szótó mozgó magánhangzója nem esett ki, vagy a köznyelvi normától eltérő kötőhanghasználat jelentkezett: [boko^jot] 2;6,12 'bokrot', [boko^joko^t] 2;6,12 'bokrokat', [e^jepes'] 2;8,1 'epres', [e^jepet] 2;6,5 'epret', [s'a:to^jot] 2;7,29 'sátrat', [s'ɔ^jɔdot:] 2;7,13 'szaladt', [pisto^jot] 2;8,2 'pisztolyt', [fy^jɔdot:] 2;7,30 'fürdött', [busot] 2;5,27 'buszt'.

2.2.3. Hangelhagyás

Ez a jelenség főleg a megfigyelés kezdeti időszakában fordult elő, s elsősorban az abszolút szóeleji és az abszolút szóvégi hangokat érintette: [mos'] 2;1,27 'most', [hide] 2;1,2 'hideg', [tati] 2;3,15 'statek = hajó', [me:] 2;7,23 'még', [maju] 2;5,24 'Maluch = Kispolski', [minda:j] 2;2,5 'mindjárt', [mej] 2;6,12 'mert', [ɔjud'a:] 2;5,18 'aludjál', [a:z] 2;0,16 'ház', [ho:j] 2;0,16 'hold', [e^jej] 2;7,9 'tenger'. Előfordult, hogy több hangból álló szóvégeket is elhagyott: [je:zu] 2;1,28 'Jézuska', [bits'i] 2;5,12 'bicikli', [mɔ], illetve [ma] 2;0,16 'maci', [babu] 2;3,1 'babunia = nagy'.

A szóbelseji szótagok valamelyikének eliminációja a későbbi időszak jellegzetessége volt, amikor a szavak toldalékolásával már hosszabb hangsorok alakultak ki, s ezek artikulációja nehézséget okozott: [betɔt:e] 2;5,9 'bekötözte, bekötötte', [kin'it:ɔm] 2;5,27 'kinyitottam', [kiɔntem] 2;6,23 'kiöntöttem', [beyt:em] 2;7,8 'beütöttem', [butɔs'a:t] 2;7,9 'butaságot', [s'e:ta:jom] 2;7,26 'sétáltatom', [tɔvet:ø] 2;8,4 'következő', [tutut:ɔm] 2;5,28 'kukucskáltam', [ejyt:e] 2;7,29 'elütötte'.

2.2.4. Mássalhangzó-rövidülés

A mássalhangzó-rövidülés ritka jelenségnek számított, csak szórványosan fordult elő, s általában alkalmi artikulációs pontatlanságnak minősíthető: [tɔzetɔ] 2;1,5 'kazetta', [e:tit] 2;1,5 'és itt', [hæjo:] 2;2,7 'halló', [it] 2;2,7 'itt', [o:t] 2;2,10 'ott'.

2.2.5. Hasonulások

G. nyelvében a hasonulások a magyar nyelv szabályainak megfelelően alakultak. A tipikusan gyermeknyelvi jelenség, a nem szom-

szédos hangok hasonulása, vagyis a távhasonulás (Kassai 1978; Gósy 1978) is a G. nyelvhasználatát jellemző hangtani sajátosságok közé tartozott. A távhasonulással létrejött alakok fonetikailag jóval differenciálatlanabbak, mint köznyelvi megfelelőjük, pl.: [mimɔmo:] 'villamos', [te:toto:] 'krémtúró' stb. Kassai adatai szerint a távhasonulás (a köznyelvi hasonulásoktól eltérően) általában előreható irányban következik be. G. anyagában azonban kb. egyenlő arányban fordulnak elő a regresszív, illetve a progresszív irányba ható változások. Példák regresszív hasonulásra: [mim:ɔmo:, mim:ɔmos'] 2;0,16., 2;1,9 'villamos', [jojod] 2;1,15 'forog', [e:tit:] 2;2,5 'és itt', [poja:j] 2;7,4 'pohár'; progresszív hasonulásra: [pepi] 2;0,26 'Peti', [te:toto:] 2;0,18 'krémtúró', [vɔninija:s'] 2;5,25 'vanília', [piŋviŋ] 2;7,28 'pingvin', [piŋviŋ ejni] 2;8,2 'pingvingelni = úgy járni a hálósákban, mint egy pingvin'. Általában a gyermek hangállományában régebben kialakult, tehát jobban megszilárdult hang indukálja a hasonulást.

2.2.6. Hiperkritikus ejtés

Több gyermeknyelvi tanulmány is beszámol arról a jelenségről (S. Meggyes 1971; Gósy 1979), hogy egy újonnan megtanult hang olyan hangok helyén is megjelenik, amelyekkel korábban a gyermek helyettesítette az adott (új) hangot. A túlgeneralizált képzés következtében "hiperkritikus" ejtés jelentkezik: [keŋej] 2;7,9 'tenger', [mokoŋ] 2;7,13 'motor', [kihoskoŋ] 2;7,8 'ki hozta?', [ja:ke:kbo:j] 2;8,4 'játékból'. G. esetében a hiperkritikus ejtés kizárólag a korábbi k > t hanghelyettesítések helyén fordult elő. Ennek az ejtéssajátságnak jellemzően nyelvpszichológiai oka van. G. abszolút biztonsággal különböztette meg a [k] és [t] hangokat, de éppen a korábbi helyettesítés miatt az adott hangsorban jelentkező mássalhangzó fonémaértékében bizonytalanná vált. Mindez azt eredményezte, hogy néha a [t] helyén is [k] jelent meg (Gósy 1984 alapján).

Összegzés

A vizsgálat eredményei azt mutatják, hogy a nyelvelsajátítás e korai fázisában a nyelvi interferencia-jelenség a beszéd fonetikai szintjén is megjelenik. G. esetében a magánhangzó-harmónia mint sajátosan ma-

gyar nyelvi szabály alkalmazását a második nyelv eltérő nyelvi-fonetikai rendszere valószínűleg befolyásolja. A gyermeknyelvi szakirodalomban alig akad példa a magánhangzó-illeszkedés elmaradására (Kenyeres 1928; S. Meggyes 1971), vagyis az egynyelvű magyar gyermekek beszédtevékenységét e sajátosság nem jellemzi. Feltehetően a két nyelv hatásának tudható be G. fonémaállományának és beszédhang-realizációjának sajátos alakulása is. Így például az egynyelvű magyar gyermekek hangkészletében viszonylag későn, a legutoljára elsajátított beszédhangok egyikeként jelenik meg a [dʒ], a vizsgált gyermek esetében ez azonban jóval korábbi, s (eleinte) ez volt az egyetlen olyan affrikáta, amely nem kizárólag helyettesítő hanggal realizálódott. A többi affrikáta ugyanakkor - az artikulációs "megoldást" tekintve - egy másik csoportba tartozott. Az elemzések alapján valószínűnek tűnik, hogy a kétnyelvű gyermek beszédének hangtani sajátosságait - mindkét használt nyelv esetében - a hasonló megoldási stratégiák jellemzik, azaz nem az egyes nyelvek fonetikai különbözősége a meghatározó.

Megjegyzés: Nyomdatechnikai okokból a gyermeknyelvi adatok átírásában a következő megfeleléseket alkalmaztuk: a veláris, felső nyelvvállású, illabiális magánhangzó=î; a köznyelvi palatális, alsó nyelvvállású, illabiális magánhangzó az *erdő* szó átírásában=ε; a felső vessző a palatalizáltságot jelzi.

Irodalom

Gósy Mária: Hangtani és szótani vizsgálatok hároméves gyermekek nyelvén. NytudÉrt. 119. Akadémiai Kiadó. Budapest 1984.

Gósy Mária: Szavak és toldalékok hangalaki jellemzői a gyermeknyelven. Magyar Fonetikai Füzetek 2. Budapest 1978.

Kassai Ilona: A hangsúly kialakulása a gyermeknyelven. Magyar Fonetikai Füzetek 7. Budapest 1981.

Kassai Ilona: Az intonáció szerepe a gyermeknyelven. Magyar Fonetikai Füzetek 3. Budapest 1979.

Kassai Ilona: Távhasznulás a gyermeknyelven. Nyelvtudományi Közlemények 83. Budapest 1981.

Kenyeres Elemér: A gyermek beszédének fejlődése. Studium Kiadó. Budapest 1928.

Lengyel Zsolt: A gyermeknyelv. Gondolat. Budapest 1981.

Leopold, W.: Speech development of a bilingual child. Northwestern University Press. Evanston 1949.

S. Meggyes Klára: Egy kétéves gyermek nyelvi rendszere. NytudÉrt 73. Akadémiai Kiadó. Budapest 1971.

A. Molnár Ildikó: A hanghelyettesítés típusai a gyermeknyelvben 18-22 hónapos kor között. Magyar Fonetikai Füzetek 1. Budapest 1978.

Montágh Imre et al.: Gyakori beszédhibák a gyermekkorban. Tankönyvkiadó. Budapest 1993.

Pavlovich, M.: Le langage enfantin. Champion. Paris 1920.

Ronjat, J.: Le development du langage observé chez un enfant bilingue. Champion. Paris 1913.

Vértes O. András: A beszédhibákról. Nyelvőr Füzetek 5. Akadémiai Kiadó. Budapest 1956.

A TRANSZFORMÁCIÓS ÉSZLELÉS FEJLŐDÉSE ÉS ZAVARAI

Gósy Mária
MTA Nyelvtudományi Intézete

Bevezetés

A beszédészlelés fejlődése az újszülöttkor időszakában kezdődik; majd fokozatosan kialakul az anyanyelv-elsajátítás folyamán az a mechanizmus, amelynek révén a gyermek anyanyelve beszédhangjait, hangkapcsolatait azonosítja, megkülönbözteti, felismeri az ismétlődő hangsorokat (Affolter 1974; Gósy 1989). A beszédészlelésnek különböző részfolyamatai vannak, amelyek fejlődése az életkori függvényében különböző. Az intonációs struktúrák azonosítása korábban válik biztossá, mint a szegmentális szint egységei; az akusztikailag nagyobb különbségek felismerése korábban bekövetkezik, mint az akusztikailag kisebb különbségeké. A beszédritmus pontos észlelése már hároméves kor körül tetten érhető, míg az egy jegyben különböző mássalhangzók elkülönítése többnyire csak hatéves korra alakul ki.

A transzformációs észlelés az észlelési szint működésének egyik részfolyamata: a gyermek képessé válik a beszédhangokat valamely vizuálisan észlelhető jelenségre (tárgyra, alakzatra, rajzolatra) átvinni és egyiket a másik alapján azonosítani. A transzformációs észlelés feltétele, hogy a gyermek képes legyen: (i) beszédhangnyi időtartamú (kb. 50-200 ms) hangjelenségek azonosítására, (ii) vizuális jelenségek felismerésére, (iii) a feltételes reflex elve alapján kialakított kapcsolat megértésére és fenntartására, továbbá (iv) mindezek ismételt működtetésére. A transzformációs észlelésnek a fentiek szerint definiált működése az alapja az alfabetikus írás elsajátításának, jelen esetben tehát a magyar nyelv írott változata megtanulásának, az olvasásnak és írásnak. Ha a transzformációs észlelés bármely feltétele nem teljesül, akkor a folyamat nem működik, s ez nyilvánvalóan az olvasás és írás elsajátításának zavarához vezet (Wyeke 1983).

E tanulmányban két kérdésre kerestünk választ. 1. Melyik az az életkor, amelyre - normális anyanyelv-elsajátítás esetén a transzfor-

mációs észlelés kialakul, illetőleg melyek a fejlődés egyes állomásai? 2. Az olvasástanulási kudarcok hátterében miként jelentkezik a transzformációs észlelés zavara?

Feltehetően módszertani nehézségek miatt a transzformációs észlelés vizsgálata új a magyar szakirodalomban, meglehetősen kevés adatot találunk erre vonatkozóan a külföldi munkákban is (vö. Gerebenné 1995). Gyermek beszédészlelési folyamatának diagnosztikus megítélésében szükséges a transzformációs észlelés tesztelése is (Gósy 1995). Ennek a módszerét alkalmaztuk jelen kísérletsorozatunkban.

A kísérleti módszer

A teszt lefolytatásának időtartam kb. 4 perc egy gyermekkel; 12 darab 2 centiméteres színes fakocka (piros, zöld, sárga és fekete, minden színből 3 kocka) szükséges hozzá. A teszt négy rövid részből áll. A gyermek elé kitesszük a 12 fakockát, s megkérjük, hogy válasszon közülük egyet. Megegyezőnk a gyermekkel, hogy a kiválasztott fakockának nevet adunk, s elnevezzük úgy, hogy "p". Ha a gyermek felfogta a kocka és a kiejtett beszédhang közötti kapcsolatot, folytathatjuk a tesztelést. I. Megkérjük a gyermeket, hogy tegye ki maga elé az asztalra azt, hogy "p, p". A gyermek feladata az általa korábban kiválasztott színű kockából kettő kirakása. II. Majd arra kérjük a gyermeket, hogy tegye ki azt, hogy "p, ó". Azt várjuk, hogy a gyermek a korábban azonosított kocka mellé egy (bármilyen) más színű kockát tesz ki. III. Ezúttal az "ó, p" kirakását kérjük a gyermektől. IV. Végül a "b, ó, p, ő" hangok kirakását kérjük. (A balról jobbra irány megtartásának segítségével minden próbában ujjunkkal balról jobbra tartó vonalat húzunk az asztalra, jelezve ezzel a gyermeknek, hogy hova és hogyan rakja ki a kockákat.) Minden próba (sikeres vagy sikertelen) elvégzése után a kockákat összekeverjük, így a gyermeknek mindig a 12 rendelkezésre állóból kell választania. A jó megoldást jelekkel szemléltetjük.

- | | |
|------|---|
| I. | <input type="checkbox"/> <input type="checkbox"/> |
| II. | <input type="checkbox"/> ■ |
| III. | ■ <input type="checkbox"/> |
| IV. | ◆ ■ <input type="checkbox"/> ☒ |

Kísérleti személyek

A kísérletet óvodás és iskolás gyermekekkel végeztük el. Az óvodások mintegy fele lány, fele fiú, 24 négyéves, 20 ötéves, 25 hatéves, 20 hétéves, összesen 89 fő; az iskolások olvasási nehézséggel küzdő tanulók: 35 hétéves, 45 nyolcéves és 35 kilencéves gyermekek, kétharmaduk fiú; összesen 115 iskolás.

Eredmények

I. A transzformációs észlelés kialakulása

A teszt egy-egy próbájának értéke helyes megoldás esetén 25%, így a teljes teszt jó eredménye 100%, ezt az értéket nevezzük 'megoldási biztonság'-nak. Milyen működéssorozatot tételeznek fel az egyes próbák. Az első a jel, jelen esetben a beszédhang és az alak, jelen esetben egy színes kocka azonosítása, illetőleg ismételt reprodukálása. E feladat sem egyszerűen a hang és a kocka megfeleltetését vizsgálja, hiszen bizonyos ritmusészlelés is szükséges a hibátlan megoldáshoz: nem egy, nem három, hanem két azonos színű kockát kell a gyermeknek kitennie. Minthogy a ritmusészlelés már hároméves kor körül kialakul, így e feladat megoldása 4 éves korra elvárható. A jel és alak megfeleltetése az anyanyelv-elsajátítás korai szakaszaiban alakul ki: a gyermek az elhangzó hangsort valamiféle szemantikai tartalommal kapcsolja össze (S. Meggyes 1971; Lengyel 1983). A jelen tesztben csupán a jel természete változik bizonyos mértékig, egy izolált, zöngétlen felpattanó zárhang.

A második próba a tulajdonképpeni transzformációs percepció. A gyermeknek rá kell jönnie, hogy a hallott új hang, az "ó" a korábban kiválasztott és "p"-nek elnevezett kockán kívül bármelyik más színű lehet. Ekkor történik meg a transzformációs művelet: a "beszédhang--színes kocka" megfeleltetést át kell vinni a "másik beszédhang--másik színes kocka" megfeleltetésre. A művelet sor voltaképpen már csaknem a kezdetektől működik az anyanyelv-elsajátítás folyamatában: a gyermek megtanulja a "/labda/=piros, kerek, gurul (stb.)" megfeleltetést, majd megtanulja az "/alma/=piros, kerek, édes, ehető (stb.)" megfeleltetést, amelyek alapján a transzformációs művelet sor működni kezd.

A harmadik próba már összetettebb feladat: mind a hangzásra, mind egy bizonyos térbeli irányra vonatkozó szeriális észlelés szükséges a jó

megoldáshoz. A hangzás időrendiségének (szerialitásának) pontos azonosítása és reprodukálása ugyancsak a beszédtanulás korábbi szakaszaiban alakul ki, mintegy 3 éves korig (vö. Gósy 1996). A balról jobbra irány tartásához a próbában segítséget kap a gyermek: a vizsgáló vonalhúzása kijelöli számára a szükséges irányt (bár ötéves kortól elvárható a balról jobbra történő haladás tartása). A szeriális észlelés alapvetően fontos a szókincs alakulása, az olvasástechnika és az írástanulás szempontjából.

A negyedik próba komplex feladat, itt még több percepciók működés együttműködése szükséges a jó megoldáshoz. A jó transzformációs és szeriális észlelésen túl szükséges a hibátlan beszédhang-azonosítás, valamint - a korábbiaknál nagyobb mértékben - a rövid idejű verbális memória. Természetesen mindezek a készségek már jóval korábban is működnek a gyermeknél; e próba valamennyinek a biztos, összehangolt, egyidejű működését vizsgálja. Az 1. táblázat tartalmazza a teljes teszt életkorspecifikus értékeit óvodás gyermekeknél. Az átlag voltaképpen a 'megoldási biztonság'.

1. táblázat: A transzformációs észlelés életkori átlagai

Életkori csoportok	Jó megoldás (%)				Átlag
	I.	II.	III.	IV.	
4 évesek	100	50	16,6	-	41
5 évesek	100	60	40	5	51,2
6 évesek	100	75	50	50	68,75
7 évesek	100	100	75	100	93,7

Mit jelentenek az életkori adatok? A legkisebbek képesek a hang és a forma azonosítására, illetőleg a teszhelyzetben az ismételt reprodukálásra (I.=100%). Az e korúak fele megoldja a transzformációs átvitel feladatát is, vagyis kimondható, hogy 4 éves korra a transzformációs észlelés részben már kialakult, illetőleg a gyermekek egy részénél kialakulóban van (II.=50%). A harmadik feladat megoldására a négyévesek 16,6%-a volt képes. Az utolsó próbát nem tudták megoldani. Az ötévesek az első teszt hibátlan megoldására képesek

(I.=100%); a hang és forma azonosításán túl a transzformációs átvitelt kissé jobban oldották meg, mint a négyévesek (II.=60%). A gyermekek 40%-a a fordított elhangzású transzformációs feladatot is megoldotta (III.), többségük azonban a szerialitás tekintetében tévedett: nem a megfelelő sorrendben rakták ki a kockákat. Az utolsó feladat megoldására az e korúak lényegében képtelenek, bár találtunk olyan gyermeket, aki megoldotta a próbát. A hatévesek - várhatóan - hibátlanul megoldják a hang és forma azonosítását (I.); szignifikánsan változik a transzformációs percepció jó működése (II.=75%), a gyermekek fele oldja meg hibátlanul a harmadik próbát (kismértékű változást látunk az előző korcsoporthoz képest), s ugyanilyen mértékben jó a teljesítményük a negyedik próbában, ez ismét szignifikáns fejlődést mutat (IV.=50%). A hatévesek teljes transzformációs észlelése jól működik: I., II. és IV. próbában 100%-os eredményt hoznak, a biztos irányismeretet is feltételező harmadik próbánál azonban adódnak még tévedések (75%).

A kísérlet során nyilvánvalóvá vált, hogy a gyermekek hibás eredménye nemegyszer a szerialitásban jelentkező problémára vezethető vissza. Ez azt jelenti, hogy a gyermek képes volt ugyan a transzformációs percepcióra (II. és III. próba), azonban az irányfelismerési bizonytalanság, illetőleg az elhangzás időrendjének és a térbeli sorozatalkotás irányának összehangolása nem volt tökéletes. Ha eltekintünk a szerialitás okozta tévedésektől, akkor az életkori teljesítményértékek jelentősen változnak. Ezeket az adatokat mutatja a 2. táblázat.

2. táblázat: A transzformációs észlelés korrigált teljesítményértékei

Életkori csoportok	Jó megoldás (%)				Átlag
	I.	II.	III.	IV.	
4 évesek	100	66,6	49,9	-	54,1
5 évesek	100	100	70	5	68,75
6 évesek	100	100	87,5	50	84,3
7 évesek	100	100	100	100	100

A szerialitási zavar a 4-6 évesek helyes megoldását 13-17%-ban, a hatévesekét közel 7%-kal csökkenti. Kimondható, hogy a hang és a

forma azonosítása és önálló reprodukálása már négyéves korban hibátlanul működik, ötéves korra pedig kialakul a transzformációs észlelés. Hatéves korban lényegesen javul a III. próba eredménye, s a gyermekek fele a komplex transzformációs feladat teljesítésére is képessé válik. A teljes teszt hibátlan megoldása, azaz a tökéletes megoldási biztonság hétéves korra várható el (bár a kísérletben részt vevő hétéves óvodások mintegy 7%-ánál szerialitási zavart találtunk).

II. A transzformációs észlelés zavarai

A) Olvasási nehézséggel küzdő gyermekek transzformációs észlelése

Láttuk, hogy az olvasás és írás elsajátítása nagymértékben a transzformációs észlelése alapul. Megnéztük, hogy miként teljesítenek e tesztben azok az iskolások, akiknek rövidebb-hosszabb ideje fennálló olvasástanulási nehézségeik vannak. A gyermekek valamennyien hét éven felüliek, tehát az elvárt teljesítményük a négy próba hibátlan megoldása, azaz 100%. A mért eredményeket a 3. táblázat tartalmazza. Az átlag ismét a megoldási biztonsággal azonos.

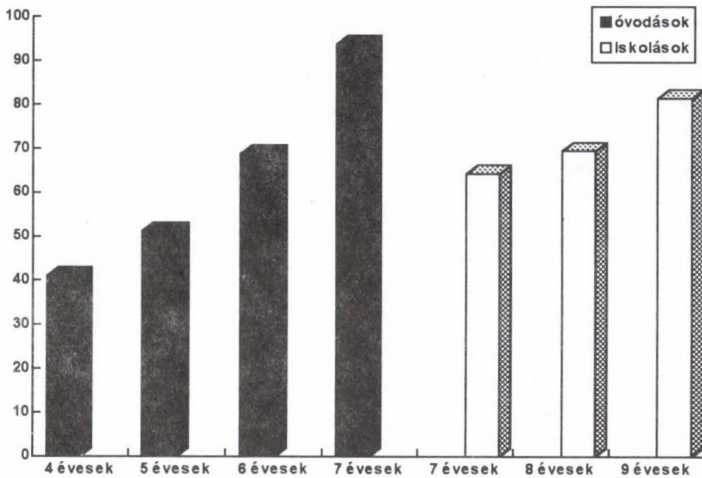
3. táblázat: Iskolások transzformációs észlelésének adatai

Próbák	Jó megoldás (%)		
	7	8	9 (év)
I.	82,8	80	97,1
II.	80	71,1	80
III.	62,8	80	85,7
IV.	31,4	46,6	62,8
Átlag	64,25	69,4	81,4

A 7 és 8 évesek az első próba megoldásában nem mutatnak szignifikáns különbséget; csaknem 100% ugyanakkor a harmadikosok teljesítménye. A második próbában egyik csoport teljesítménye sem mutat javulást, ami azt jelenti, hogy ezeknek az olvasási nehézséggel küzdő iskolásoknak a transzformációs észlelése csak mintegy 70-80%-ban működik. Tekintetbe véve a normál életkori átlagokat, ez azt jelenti, hogy többéves a fejlődési elmaradásuk. Nagymértékű a zavar a harma-

dik próbában is: bár a másodikosok és a harmadikosok 80% körüli értéket mutatnak átlagosan, ez jelentős elmaradást mutat az életkori normákat figyelembe véve. A legnagyobb mértékű a zavar, illetőleg az elmaradás a negyedik próbában. E komplex feladat megoldására a tesztelt iskolások csak kis hányada volt képes, ami azért rendkívül meglepő, hiszen ezek a gyermekek - még, ha küzdenek is olvasási nehézséggel - lényegében olvasnak már és írnak.

Minden korcsoportban jelentős az elmaradás: az első próbában a gyermekek nem érik el a 4 évesek (!) teljesítményét, a másodikban a hatévesek teljesítményét közelítik, a harmadikban mintegy egy-két év elmaradást mutatnak a biológiai életkorukhoz képest. A negyedik próba eredménye szinte hihetetlen: iskolásaink a hatéves óvodások szintjét nem érik el (vö. 1. ábra).



1. ábra

A transzformációs észlelés átlageredményének alakulása óvodás és olvasási nehézséggel küzdő iskolás gyermekeknél

Kísérletünkben 9 évesek voltak a legidősebb gyermekek, azonban néhány 10-15 éves iskolás gyermek transzformációs észlelésének adatait is elemeztük. Valamennyien olvasási nehézséggel és helyesírási gondokkal küzdöttek (évek óta). Osztályfokukat tekintve harmadik, negyedik, ötödik, hatodik, hetedik és nyolcadik osztályosok (döntő többségük fiú). A feltételezésünk - hasonlóan a korábbiakhoz - az volt, hogy a súlyos olvasási nehézség hátterében a transzformációs észlelés valamilyen mértékű működési zavara meg kell, hogy jelenjen. A gyermekek egyéni adatait a 4. táblázatban összegeztük.

4. táblázat: 10-15 éves, tanulási nehézséggel küzdő iskolások transzformációs észlelésének adatai

Iskolások (életkor)	Jó megoldás (%)				Átlag
	I.	II.	III.	IV.	
10;0	-	-	-	-	0
10;2	100	-	100	-	50
10;6	100	100	100	-	75
10;7	100	100	-	-	50
10;8	100	100	-	-	50
11;2	100	100	100	-	75
11;6	100	100	100	-	75
11;8	100	100	100	-	75
12;4	-	-	100	100	50
12;6	100	-	100	-	50
13;1	100	-	100	100	75
15;3	100	-	100	-	50

A tanulási nehézséggel küzdő iskolások egy részénél a transzformációs észlelés súlyos zavarokat mutat; teljesítményük az egyes próbákban rendkívül változó, némelyikük a négyévesek eredményét sem tudja produkálni. Jellemzően a komplex működést igénylő negyedik próba jelent megoldhatatlan feladatot a számukra. Egyéni megoldási biztonságuk 0% és 75% közötti. A többségük által elért 50% és 75% az öt- és hatévesek teljesítményének felel meg.

B) Rizikó gyermekek transzformációs észlelése

1. Az óvodásokkal és az olvasási nehézséggel küzdő iskolásokkal végzett vizsgálatok alapján, kísérletet folytattunk olyan 5 és 6 éves gyermekekkel, akiknél a beszédészlelési és/vagy a beszédmegértési mechanizmus valamely részfolyamatában vagy részfolyamataiban zavart, elmaradást találtunk. A kísérletben 14 ötéves és 25 hatéves gyermek vett részt (nagyobb részük fiú). A 14 ötéves körül 5-nél a teszt egyáltalán nem volt elvégezhető: a gyermek képtelen volt akár csak az első próba elvégzésére is. Ezért a továbbiakban csak annak a 9 gyermeknek az adataival foglalkozunk, akik valamilyen mértékben képesek voltak az együttműködésre. Hipotézisünk az volt, hogy a transzformációs észlelés érintett azokban az esetekben, ahol más zavar is tetten érhető a feldolgozó rendszerben; ezek a gyermekek tehát rizikó gyermekek ennek a részfolyamatnak a jó működését tekintve. Az eredményeket a 5. táblázat tartalmazza.

5. táblázat: 5 és 6 éves rizikó gyermekek transzformációs észlelésének adatai

Próbák	Jó megoldás (%)	
	5 évesek	6 évesek
I.	66,6	76
II.	55,5	68
III.	44,4	48
IV.	-	8
Átlag	41,6	50

Az ötévesek feldolgozott adatai nem, illetőleg kismértékű eltérést mutatnak az életkori átlagoktól. Jelentős az elmaradás az első próbában, minimális az eltérés a másodikban, a harmadik nem mutat eltérést az átlagos ötévesek teljesítményétől, a negyedik ismét kismértékű elmaradást jelez. A végeredményt tekintve azonban szignifikáns különbséget látunk: a rizikó gyermekek teljesítménye a négyévesekével azonos, az elmaradás tehát a biológiai életkor szerint egy év. Hangsúlyoznunk kell, hogy ezek az adatok bizonyos értelemben "korrigáltak", hiszen annak az öt gyermeknek az adatait nem vettük figyelembe, akik a teszt elvég-

zésére képtelenek voltak. A rizikó ötéves gyermekek átlaga a közölnél valójában lényegesen gyengébb.

Az átlagot tekintve közel 20%-kal elmaradnak az életkorukban elvárt eredménytől. Az ötévesekhez képest ugyan kétségkívül fejlődés látható, de ez nem a kívánt mértékű. A rizikó hatévesek egyetlen próbában sem tudják hozni az életkori szintet.

2. A vizsgált rizikó gyermekek másik csoportja az ún. megkésebb beszédfejlődésűek. Ezek a gyermekek hároméves koruk körül kezdtek beszélni (ép hallás és ép intelligencia mellett). Tekintettel arra, hogy a beszédpercepció mechanizmusuk kisebb-nagyobb mértékben mutat zavarokat, feltevésünk ismét az volt, hogy a transzformációs észlelésük is érintett. A kísérletben 10 hatéves (nagy csoportos óvoda) és 14 hétéves és 14 nyolcéves gyermek vett részt (nagy arányban fiúk). (Az iskolások olvasási nehézséggel küzdő gyermekek voltak.) Adataikat a 6. táblázat mutatja.

6. táblázat: Megkésett beszédfejlődésű gyermekek kísérleti adatai

Próba	Jó megoldás (%)		
	6 évesek	7 évesek	8 évesek
I.	60	71,4	85,7
II.	40	85,7	85,7
III.	20	42,8	64,2
IV.	-	14,3	42,8
Átlag	30	53,5	69,6

Az eredmények önmagukért beszélnek, a feltételezésünk igazolódott: a megkésett beszédfejlődésű gyermekek transzformációs észlelése erősen zavart. Valamennyi próbában elmaradást mutatnak az életkorukhoz képest, különösen jelentős mértékben a komplex feladatmegoldást igénylő negyedik próbában. A négy próba megoldási biztonságát tekintve, a megkésett beszédfejlődésű hatévesek nem érik el a négyéves gyermekek szintjét, a 7 évesek az ötéves óvodások teljesítményét mutatják, a nyolcévesek pedig a hatévesekét. Az elmaradás átlagosan két év. (Vö. Tallal–Piercy 1977.)

III. A transzformációs észlelés működési hibái

Az egyes próbáknak megfelelően osztályoztuk a hibákat (függetlenül a kísérletben részt vevő gyermekek életkorától). Nem soroljuk fel minden esetben, ha a gyermek semmit nem tesz ki a próbában; ez minden tesztnél előfordult.

I. A jel/alak megfeleltetési nehézség többféleképpen jelentkezik a gyermekek megoldási kísérleteiben. Ha a gyermek képtelen a feladat megoldására, akkor vagy nem rak ki semmit, vagy kirak egy vagy két kockát, amelyek azonban különböznek a korábban kiválasztott és megnevezett kockától.

A próba megoldásához szükséges folyamat zavara leggyakrabban abban jelenik meg, hogy a gyermek a rendelkezésére álló mindhárom azonos színű kockát kirakja, vagyis a "p, p" helyett "p, p, p" sorozatot jelenít meg. Ritkábban fordul elő, hogy csak egyetlen kockát tesz ki, tehát a "p, p" helyett annyit, hogy "p".

Ebben a próbában nem tekintettük hibának, ha a gyermek a két kockát nem egymás mellé tette ki, hanem egymás fölé (vagy alá). Mindazonáltal ez az elhelyezés nem tökéletes, s a legtöbb esetben a szerialitás zavarára utal (ez gyakran aztán a III. próbában is jelentkezik). A megoldást azért fogadtuk el mégis jónak, hiszen a jel/alak megfeleltetés elvégzésére a gyermek képes volt.

II. A szorosan vett transzformációs észlelés zavarának is különféle megjelenési formái vannak.

a) A gyermek egy, több vagy akár sok kockát tesz ki egymás mellé, függetlenül azok színétől: sem a transzformációs észlelés, sem a jel/alak megfeleltetés nem működik. Érdekes, hogy ez a megoldás olyankor is előfordul, amikor a gyermek az első próbát jól oldotta meg.

b) A gyermek az elnevezett kockát teszi ki "p", itt megáll: a transzformáció nem működik.

c) A gyermek megismétli az első próbát, s a "p, ó" helyett újra a "p, p" sorozatot jeleníti meg: a transzformációs észlelés működése helyett tanulási effektus jelentkezik.

d) A gyermek a "p"-nek megfelelő színű kockából mindhármat kiteszi: a transzformáció hiánya ritmusészlelési zavarral társul.

f) A gyermek a "p, ó" helyett "ó, p" sort rak ki: a transzformációs észlelés ugyan működik, de a szerialításban zavar van.

g) A "p, ó" helyett "p, p, ó" jelenik meg: a transzformációs észlelés itt is működik, de a ritmusészlelési nehézség bizonytalanra teszi a gyermek döntését, s ismét a tanulási effektus győzedelmeskedik.

III. A transzformációs észlelés és a szerialítás biztos, összehangolt működése még iskoláskorban is sok nehézséget okoz.

a) A gyermek kockákat tesz ki egymás után; ez gyakran akkor következik be, ha a második próbában is tévedett.

b) Az előzőeknél összetettebb feladat megoldására a gyermek képtelen, a ritmusérzék működésének eredménye, hogy vagy a "p"-nek vagy az "ó"-nak megfelelő kockából tesz ki kettőt. Érdekes módon ez a megoldás akkor is előfordulhat, ha a második próbát a gyermek jól teljesítette.

c) A transzformációs észlelés és a szerialítás is működik, azonban ritmusészlelési zavar miatt a gyermek a kívánt "ó, p" helyett "ó, ó, p" sorozatot rak ki.

d) Az "ó, p" sorozat helyett megjelenő "p, ó" a szerialitás egyértelmű hibájára vezethető vissza. Jellegzetes, hogy a gyermekek egy része valójában jól oldja meg a feladatot: elsőként az "ó"-nak, majd a "p"-nek megfelelő kockát rakja ki, csak hogy jobbról balra. Ennek következtében az asztalon megjelenő sorozat "p, ó" lesz, ezáltal a próba megoldása nem elfogadható.

IV. A mennyiségi mutatókból nyilvánvaló, hogy a legtöbb hiba és megoldási variáció a negyedik próbában tapasztalható, annak komplexitásából adódóan.

a) A feladatmegoldás teljes sikertelenségét mutatja, amikor a gyermek kockákat tesz ki egymás után. Érdekes módon ez akkor is előfordul, ha az első három próbát jól teljesítette.

b) A tanulási effektus, illetőleg az egyszerű jel/alak megfeleltetést végzi el a gyermek a "p, p" sorozat kirakásával (ez rendszerint akkor következik be, ha az előző próbákban is hibázott).

c) A négy hang emlékezeti megtartása sok gyermeknek okoz nehézséget, illetőleg a négy hangnak megfelelő transzformáció elvégzése. Ennek a megjelenési formája a három kocka kirakása. Gyakori az

elhangzott első három hang transzformációjának elvégzése, az eredmény: "b, ó, p". A mássalhangzók azonosításának, ill. megkülönböztetésének zavara eredményezi a két zárhang lokalizációjának hibáját: "p, ó, b". A zöngés és zöngétlen zárhang felismerési zavara eredményezi az azonos mássalhangzóknak megfelelő kockák megjelenését: "p, ó, p". Előfordul, hogy egy mássalhangzó és a két kívánt magánhangzó jelenik meg: "b, ó, ö" vagy két mássalhangzó és az "ö"-nek megfelelő kocka: "b, p, ö".

d) A négy beszédhangnak megfelelő kockasorozatok kirakása a helyes ritmusészlelés működési eredménye; a többi részfolyamat esetleges zavara azonban itt is számos tévedési variációt hoz létre. Sokszor még a magánhangzók és a mássalhangzók transzformációja is hibás: "b, ö, p, b". A hangazonosítás/hangdifferenciálás zavarának következtében felépő sorozatok: "p, ó, b, ö", "p, ó, p, ó", "p, ó, p, ö" stb.

A hibatípusok elemzése rávilágít a folyamatok belső működésére, illetőleg a tévedések közötti különbségekre. Bár a próba elfogadása vagy nem elfogadása egyértelműen a helyes vagy nem helyes megoldás függvénye; a gyermek transzformációs percepciójának értékelésekor a hibák minőségi elemzése is fontos. A korrekció szempontjából lényeges annak ismerete, hogy nagymértékű zavarról vagy csak egy-egy részfolyamatot érintő hibás működésről van-e szó.

Következtetések

A transzformációs észlelés kialakulásának, fejlődésének, illetőleg zavarainak kísérletes vizsgálata a magyar szakirodalomban újnak számít. Megállapíthatjuk, hogy a négy részből álló teszt első próbáját már a négyévesek is 100%-os biztonsággal oldják meg, míg a többi próba hibátlan megoldása öt-, hat- és hétéves korban következik be. Az anyanyelv-elsajátítás számos más folyamatához hasonlóan a transzformációs észlelés kialakulása és fejlődése is életkorspecifikus.

Az adott életkornak nem megfelelő transzformációs észlelés rendszerint egyéb működési zavarokkal jár együtt, s vezet az iskolai kudarcok kialakulásához. Az eredmények azt mutatják, hogy a transzformációs észlelés zavara hozzájárul az olvasás- és írástanulási nehézségek kialakulásához, s ily módon a tanulási zavar szindrómájának

egyik forrása. Az adatok megerősítették azt a feltételezésünket, hogy a megkésett beszédfejlődés rizikó faktor - egyebek mellett - a transzformációs észlelés ép működésére nézve. Hasonlóképpen rizikó gyermekek e tekintetben azok, akiknek a beszédpercepciós mechanizmusában egy vagy több folyamat nem életkori szinten működik. A transzformációs észlelés fejlesztése - amennyiben szükséges - nagyon fontos, hiszen számos következményes probléma kialakulása előzhető ezáltal meg.

Irodalom

Affolter, F.: Leistungsprofile wahrnehmungsgestörter Kinder. Pad. Fortbildung. Praxis. Vol. 40. Basel 1974.

Gerebenné Várbíró Katalin (szerk.): Fejlődési diszfázia. Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola. Budapest 1995.

Gósy Mária: Beszédészlelés. MTA Nyelvtudományi Intézete. Budapest 1989.

Gósy Mária: GMP-diagnosztika. Nikol. Budapest 1995.

Gósy Mária: A szeriális észlelés fejlődése és zavarai. In: Gósy Mária (szerk.): Gyermekkori beszédészlelési és beszédmegértési zavarok. Nikol. Budapest 1996, 83-100.

Lengyel Zsolt: A gyermeknyelv. Gondolat. Budapest 1983.

S. Meggyes Klára: Egy kétéves gyermek nyelvi rendszere. Akadémiai Kiadó. Budapest 1971.

Tallal, P.--Piercy, M.: Defects of auditory perception in children with developmental dysphasia. In: Wyeke, M. (ed.): Developmental Dysphasia. Academic Press. London, New York 1977, 63-83.

A GÉGE NÉLKÜLI BESZÉD FONETIKAI SAJÁTOSSÁGAI

Balázs Boglárka–Gósy Mária–Szabó Iván
Szent János Kórház – MTA Nyelvtudományi Intézete

Bevezetés

Az utóbbi évtizedben ugrásszerűen megemelkedett a gégerákos megbetegedések száma. Sajnos, az esetek egyharmad részében a daganat elhelyezkedése és kiterjedése miatt az egész gégét el kell távolítani ahhoz, hogy a beteget tumorentessé tehessek. A gége elvesztése egyben a beszéd lehetőségének a megszűnését is jelenti, amely egy ember életében rendkívüli mértékű hátrányt jelent. Míg a harminc évvel ezelőtti statisztikák 10 gégerákos betegből egy nőt és kilenc férfit találtak, ma ez az arány 3:7-hez, sőt 4:6-hoz alakult. Az életkori megoszlás is változott, mert harminc évvel ezelőtt a gégerákos betegek zöme 60 év feletti volt, ma nemritkán 28, 30, 35 éves betegek kerülnek műtetre. Az átlagéletkor 60-ról 45-re csökkent. Az elmondottakból látható, hogy megváltozott a betegek összetétele; ebből következően a beszéd elvesztésének aktuális következményei is módosultak.

A gége műtéti eltávolításával megszűnik a hangadás, a beszéd megszokott, begyakorlott módja. A laryngectomy - a gége teljes eltávolítása - sok beteg számára nehezen megszokható, lényeges változást idéz elő a légutak és az emésztőtraktus viszonyában. Műtét után a száj, garat, nyelőcső, gyomor a légutaktól különálló és teljesen elzárt csőszerű egységgé alakul, ahol a lenyelt vagy szájon keresztül leszívott levegőnek csak egyetlen kitérő útja van az orr-garat felé.

Történeti áttekintés

Az első világháború után Stern (1924) - vizsgálatai alapján - úgy vélte, hogy jól beszélő betegek a gégenélküli beszédhez szükséges levegőt a gyomorban tartalékolják. Igen nagy szerepet tulajdonított a gyomorlégőanyag nagyságának és változókészségének. Seeman (1926) az özofáguszban és nem a gyomorban észlelte a levegő összegyűjtését. Már Stern is leírta, hogy néhány régóta beszélő betegnek levegő-

tartaléka spontán a gyomorból feljebb, az özofáguszba vándorolt. Dénes (1938) azon megfigyelését közölte, hogy több év óta póthangon beszélő betegének beszédjellege megváltozott, mert beszéd közben nem nyeléssel, hanem szívással pótolja a gégenélküli beszédhez szükséges levegőmennyiséget. Beckmann (1954) közleményében leszögezi, hogy annál folyékonyabb és összefüggőbb a laryngectomizált beteg beszéde, minél közelebb esik az özofáguszban lévő légzsák a póthangréshez és minél mozgékonyabb ez a "pszeudoglottisz".

Hazánkban Sáfrán Antal foglalkozott behatóan a nyelőcsőbeszéd sajátosságaival. Kimutatta és saját maga készítette műszerrel mérni is tudja a tápcsatornában uralkodó - és így a nyelőcsőbemenetre vonatkozó - nyomást. Ennek alapján vélelmezni lehet, hogy a beteg milyen könnyen vagy nehezen tud megtanulni beszélni, illetve képes-e egyáltalán nyelőcsőhangot létrehozni (Sáfrán 1981). Összehasonlította a természetes beszéd, a hangprotézissel történő beszéd és az özofágusz beszéd hangszínekeit, és megállapította, hogy nyelőcsőhang kissé jobb minőségű (a jól beszélő laryngectomizáltaknál), mint a hangprotézissel képzett beszéd (Sáfrán 1988).

A nyelőcsőhang képzése

A gégenélküli beszéd alapja a nyelőcsőhang-képzés. A név Seemantól (1926) származik. Nyelőcsőhang-képzéskor az özofágusz középső és felső része alkotja a légáramlást tápláló új légzsákot, a nyelőcsőbemenet pedig a póthangszalag - a m. laryngopharyngeus felsőhátsó részének párnaszerű kidomborodása fonációkor - kialakulásával póthangrést alkot. A beszéd tehát a póthangrésen formálódott nyelőcsőhang artikulációjával jön létre.

A nyelőcső fonációs tevékenységét az hozza létre, hogy fala a légzőmozgásokkal és az intratorakális nyomásingadozással összefüggően mozog. Ezen kívül pedig az is megkülönbözteti a nyelőcső többi részétől, hogy ezen a területen harántcsíkolt izomzat található, amely akaratlagosan működtethető. A nyelőcső lumene belégzéskor kissé kitágul, kilégzéskor beszűkül. Ezt Seeman (i.m.) egy maga szerkesztette nyelőcsőszonda segítségével grafikusán is regisztrálta és a fent leírtak magyarázatát az özofágusz anatómiai adottságaiban találta meg. A gége

motoros beidegzését ellátó nervus recurrens a nyelőcső falába is ad ágakat. A nyelőcsőben lejátszódó fonációs jelenségeket a nyelőcső-izomzatba kisugárzó hangképző impulzusokkal lehet magyarázni.

A nyelőcsőhang-képzés elsajátítása

A nyelőcsőhang beszéd taníthatósága, tanulási folyamatai, alkalmazásai, illetőleg a nehézségek ismertek (lehetőleg még a műtét előtt ismertetni kell a beteggel a nyelőcsőhang-képzésnek és a vele kapcsolatos beszéd kialakulásának elméletét és a várható eredményt). A tapasztalat szerint az operált betegek egyharmada igen jól megtanul beszélni, egyharmada szavakat, rövid mondatokat érthetően mond, egyharmaduk viszont nem tudja ezt az "artikulációs folyamatot" elsajátítani.

A kifogástalan gégenélküli beszéd egyik sarkpontja, hogy a beszéd-légzés független legyen a vitális légzéstől (Kenessey 1957). A függetlenítést nem mindig könnyű megérteni, főleg létrehozni, ha a beteg idő előtt túlzottan gyorsan igyekszik beszélni. A hangképzési gyakorlatok a levegőnyelés megtanulásával kezdődnek: nyeléssel levegőt juttatunk a nyelőcsőbe. Ezt a levegőt haspréshez vagy bőfögéshez hasonló mozdulat segítségével azonnal kinyomjuk, igyekezve, hogy a nyelőcsőhang létrejötte kísérje a folyamatot. A "póthang", pontosabban "pótzöngé" lényegesen gyengébb a normális zöngénél, mivel a nyelőcső kapacitása is lényegesen kevesebb a tüdő kapacitásánál. Sokkal a mélyebb a hangszalaghangnál, kb. 50 Hz, mert vastkosabb, durvább felületek rezgéséből keletkezik.

A műtét utáni beszédrehabilitációnak a laryngectomizált beteg életében óriási a jelentősége. Noha a nyelőcsőhangot használó betegek beszéde minőségileg alacsonyabb rendű póthang segítségével jön létre, sokan kiemelkedő beszédteljesítményre lesznek képesek. Az új hangképzés és a beszédmechanizmus megtanulása nagyon nehéz feladat. A beszéd-légzés és a vitális légzés kettéválasztása, a frissen operált és fájdalmas terület használatával lelkileg is nagyon megterheli a beteget. A beszéd elsajátításának időtartama a beteg egyéni ügyességétől, a helyi anatómiai viszonyoktól és a tanár lelkiismeretességétől függ. A laryngectomizált betegek jelentős része még aktív dolgozó. Bármilyen munka vállalása ugyanakkor legalább minimális beszédet szükségessé tesz.

Amennyiben a beteg nem tudja megtanulni a nyelőcsőhangbeszédet, akkor külső hangprotézis használata válik szükségessé. Ez a készülék elektromágneses úton, membránrezgés által kelt hangot. Ez a hang pótolja a hangszalagrezgés hiányát. A beteg a készüléket az áll alatti lágyrészterületre nyomja; a bőr és az izomzat a szájba és a garatba vezeti a rezgéseket. Az ily módon létrehozott beszéd közben a beteg némán artikulál. A Servox használata az egyik kéz szabadon tartását követeli meg, tehát csak egyes munkakörökben használható. (Az újabban külföldön egyre elterjedtebb és hazánkban is néhány esetben már jól sikerült, de rendkívül költséges belső hangprotézis, a Bloom-Singer protézis használatakor a beszélőnek egyik kezével be kell fognia a sipoly-nyílást, hogy beszélni tudjon.)

Frint a nyelőcsőbeszéd minőségének jellemzésére a beszédtanulás folyamán négy stádiumot különböztet meg (1982):

- a) a beteg képessé válik a nyelőcsőhang létrehozására,
- b) egy, két szótagból álló szavakat ejt,
- c) rövid mondatokat képez,
- d) folyamatosan beszél.

Jelen tanulmányunkban az oesophagus beszéd akusztikai sajátosságait vizsgáljuk a tekintetben, hogy az objektív adatok mennyire alkalmasak a nyelőcsőhang-beszéd minősítésére. A cél annak a feltárása, hogy mik a nyelőcsőbeszéd szubjektív minősítésének és az artikuláció objektív akusztikai következményeinek az összefüggései.

Kísérleti anyag és módszer

Kísérletünkhöz a Szent János Kórházban 1995-ben laryngectomián átesett és a nyelőcső-beszédet elsajátított betegek közül nyolc 45 és 65 év közötti férfit választottunk ki. Négyen aktív dolgozók maradtak a műtétet követően is, ketten már korábban nyugdíjasok voltak. (Iskolázottságuk különböző volt.) A betegeket a foniáter beszédük szubjektív véleményezése alapján választotta ki; voltak jól, közepesen és kevésbé jól beszélők. A kísérleti beszédanyag izolált szavakból (a hét napjai és a hónapok nevei) állt (az olvasás kiiktatása érdekében), valamint rövid, egyszeri hallásra megjegyezhető 10 mondatból állt. A mondatokat a kísérleti személyeknek az elhangzást követően azonnal

ismételniük kellett. Egy ötvenéves, átlagos alaphangmagasságon beszélő férfi rögzített hanganyaga képezte az ép beszéd referenciáját.

A felvett beszédanyag akusztikai szerkezetét az MTA Fonetikai Laboratóriumában a CSL 50-es digitális jelfeldogozóval (különböző beállításokban) elemeztük. Vizsgáltuk a frekvencia- és az intenzitás-szerkezetet, az időviszonyok alakulását, a felhangtartományt, az egyes beszédhangok akusztikai szerkezetét.

A 8 beszélő hanganyagát - ilyen beszédhez nem szokott - öt egyetemi hallgatóval hallgattattuk le. Két feladatot kellett elvégezniük. a) az első a szavak és mondatok érthetőségének megítélése volt; b) a második a beszédminőség szubjektív értékelése. 'Beszédminőségen' a hallgatóra gyakorolt esztétikai hatást értjük, amely relative független az érthetőségtől. A magyar iskolaarendszerben használatos öt érdemjegy alkalmazását kértük: 5-sel a jól érthető, illetőleg a jó beszédminőségű; 1-sel a nem vagy alig-alig érthető, illetőleg gyenge minőségű beszédet minősítették.

Eredmények

1. Akusztikai elemzési adatok

A nyelőcsőhangbeszéd rezgésképeinek periodicitása - várhatóan - eltér a normál zöngéről készült regisztrátumokon látható képtől. Az eltérés mértéke azonban az egyes beszélők esetében jellemzően különböző. Némely esetben a rezgéskép mutat periodikus szakaszokat, más esetekben gyakorlatilag egyáltalán nem találhatók benne ilyen részletek. A részben periodikus rezgésképe a zöngéhez lényegesen közelebbi hangzás benyomását eredményezi, mint a zöngés szakaszokat nem tartalmazó, teljesen aperiodikus rezgésképe. Minél ritmikusabban, tehát közel azonos időszerkezeti formációban fordulnak elő a nyelőcsőhangbeszéd periodikus szakaszai, annál természetesebb a beszéd hangzása; annál jobb lesz a beszéd szubjektív minősítése.

A felhangtartomány elemzési adatai is jellegzetes eltérést mutatnak nem csupán az ép beszéd és a nyelőcsőhangbeszéd akusztikai megjelenése, hanem az egyes nyelőcsőhangbeszéddel beszélők között is. A felhangtartományt 1200 Hz-ig vizsgáltuk. Azt tapasztaltuk, hogy míg a normál beszéd esetében a zöngé felhangjai intenzíven kimutathatók 400-

500 Hz-ig, az adott beszédhangoktól függően, addig a nyelőcsőbeszéd esetén lényegesen kevesebb intenzív felhang jelenik meg. A zönge hiányában az aperiodikus rezgések eredménye a zörejszerű akusztikai szerkezet, amely azonban nem korlátozódik a 400, 500 Hz-es tartományra, hanem az elemzett 1200 Hz-es tartományon belül végig jelentkezik. A zörejszerű akusztikai jelenségek intenzitása kismértékben változó, jellegzetes frekvenciaeltéréseket nem mutat; időbeli eltéréseket azonban látunk. Ez egyszersmind az oesophagus-beszéd szaggyatottságából is adódik. Megnéztük, milyen akusztikai eltéréseket látunk az 1200 Hz-es felhangtartomány területén a jól érthető és a kevésbé érthető beszéd között (vö. 1. táblázat).

1. táblázat: A jól és a nehezen érthető nyelőcsőhang-beszéd akusztikai sajátosságai 100-1200 Hz-es tartományban

Paraméterek	Jellemzés	
	jól érthető beszéd	alig érthető beszéd
<i>frekvencia-szerkezet</i> →	zörejes a teljes spektrumban frekvenciák szerint változó	zörejes a teljes spektrumban frekvenciák szerint nem változó
<i>intenzív felhangok</i> →	1-2 beszédhangonként	nem kimutatható
<i>intenzitás-szerkezet</i> →	frekvenciafüggő 200-400 Hz között intenzív	független a frekvenciától 200-400 Hz között nem intenzív
<i>időszereket</i> →	szegmentumkövető	nem szegmentumkövető

A beszéd teljes intenzitás szerkezetének elemzése egyúttal a nyelőcsőbeszédhez nélkülözhetetlen levegőkontroll sikerességét is jelzi. A normál beszéd intenzitás burkológörbéjére az jellemző, hogy a közlés kezdetétől közel-fokozatosan csökken az intenzitásérték a közlés végére. Noha az abszolút értékeket számos tényező befolyásolja, a különbség átlagosan 30 dB. Az özofágusz beszéd intenzitás-burkológörbéje nem mutatja a közel-egyenletesen csökkenő lefutást. Az intenzitás szerkezet alakulása ugyanakkor jellemzően különbözik a jól és a nem jól beszélők esetében.

A jól beszélőknél kúp alakú lefutást látunk: az intenzitás kezdő értéke a közlés folyamán növekszik, majd közel-azonos értéken marad, s a közlés végefelé a kezdethez hasonló értékre csökken. A kezdő/vég és a közép abszolút értékeinek különbsége átlagosan 25 dB, a szórás 15-30 dB. A közlés közepén az intenzitáscsúcsok eltérései akár 30 dB-t is elérhetnek, ez az érzelemmentes, normál beszédben nem fordul elő. A nem vagy alig érthető beszéd intenzitás-burkológörbéje fogaskerékre emlékeztető vonulatot mutat. Ezen belül a kezdő és a befejező szakaszok intenzitása kissé magasabb a közlés közepének intenzitásánál. Az eltérés átlagosan 30 dB. Ez az intenzitásstruktúra szinte "ellentettje" a jól beszélőknél tapasztaltaknak; s nagyobb mértékben tér el a normál beszédben megszokottól. A közlés egyes intenzitáscsúcsai között jelentős értékbeli különbségek vannak, amelyek a 20 dB-től egészen 60 dB-ig terjednek. A jól és gyengén beszélő betegek intenzitás szerkezetének fontos eltérése még az utóbbiak túlsúlyra emlékeztető intenzitáscsúcsai, az előbbieké jóval kerekítettebb, tehát időben lassabban lefutó intenzitásváltozásaihoz képest. Ez utóbbi ismét a normál beszéd intenzitásalakulását közelíti meg.

A nyelőcsőhangbeszédet használók feltűnően lassabban beszélnek, mint az ép emberek. Ez magától értetődően adódik a beszéd képzéséből. Az azonos közlések összehasonlító elemzéséből kitűnik, hogy az özofágusz-beszédet jól használók beszédtempója mintegy másfélszer lassúbb az ép beszédűeknél, míg a rosszul beszélőké nemritkán kétszer lassúbb. Az egyes szavak artikulációja az őket felépítő beszédhangoktól függően eltéréseket mutat. Például egy mondat teljes időtartama ép beszéd esetén 3,2 másodperc volt; egy a nyelőcsőbeszéddel jól beszélő adatközlőé 5,4 másodperc, egy gyengén beszélő adatközlőé pedig 7,7 másodperc (2. táblázat).

A magánhangzók képzési időtartama a jól beszélő laryngectomizáltaknál nagy hasonlóságot mutat az épekéhez, míg a gyengébben beszélőknél ez nem áll. A mássalhangzók időtartamának realizálása erősen függ a mássalhangzó típusától valamennyi beteg beszédében.

A zöngé gerjesztésű (ép) és a nyelőcsőhang gerjesztette özofágusz beszéd invariáns elemeinek elemzése sajátos eredményt hozott. Az özofágusz beszéd akusztikai szerkezetét döntően meghatározza a nagy

mennyiségű zörejelem. Ezek a zörejek a beszédet fonetikailag tekintve funkciótlanok, a nyelőcsőhang gerjesztés képzési eredményei. E zörejek akusztikailag a teljes beszédsspektrumban jelentkeznek, nagyjából azonos intenzitással mintegy a 100-8000 Hz-es tartományban találhatók. Ezek a zörejelemek megnehezítik a nyelőcsőhangbeszéd akusztikai szerkezetének a pontos elemzését, illetőleg adatolását.

2. táblázat: Beszédtempó-különbségek

Beszélők	Beszédtempó (hang/s)	
	átlag	szórás
ép beszélő	12	9-14
nyelőcsőbeszéd- del beszélők	7	3-10
jól érthetők	8	6-10
gyengén érthetők	6	3-8

A magánhangzók invariáns elemei, a formánsok - számukat tekintve korlátozottan ugyan -, de megjelennek. A normál beszédben rendszerint 3-4 formánst különítünk el, az oesophagus beszédben 1-2, ritkán 3 formáns azonosítása végezhető el. Természetesen a meghatározható formánsok száma nagymértékben előjelzi az adott magánhangzó felismerhetőségét, ezáltal a szó, illetőleg a mondat érthetőségét. Vizsgálatainkban azt találtuk, hogy a jól beszélő betegek beszédében 2-3 formáns jelenik meg, míg a gyengén beszélőkében általában csak egy. A zörejelemek sok esetben elfedik a zörejes mássalhangzókat, így azok elkülönítése meglehetősen nehéz, még a jól beszélő betegek esetében is. A laryngectomizált betegeink magánhangzóinak első formánsai mintegy 20%-kal magasabb frekvenciaértékeken realizálódnak, mint a hasonló alaphangmagasságú zöngével beszélő személynél. (Az 1. és a 2. ábrán egy jól és egy gyengén beszélő laryngectomizált mondatáról készült hangszínekp látható.)

2. *Percepció*s adatok

A szubjektív beszédminősítés kísérleti adatait a 3. táblázatban összegeztük. (A minősítésben használt érdemjegyeket százalékosan fejeztük ki: 1=0%, 5=100%.)

3. táblázat: Laryngectomizáltak szubjektív beszédminősítése

Beszélők	Szubjektív beszédminősítés (%)		
	átlag	szórás	érdemjegyközelítés
1. beszélő	75,2	50-100	4
2. beszélő	73,8	50-100	4
3. beszélő	68,7	50-75	3-4
4. beszélő	56,3	25-75	3
5. beszélő	37,7	25-50	2-3
6. beszélő	35,1	25-50	2-3
7. beszélő	31,8	25-50	2
8. beszélő	29,5	0-50	2

A szórásértékek sávja azt mutatja, hogy a nyelőcsőbeszédet hallgatók szubjektív ítélete kevésbé különbözött egymástól; a legjobbnak és a leggyengébbnek ítélt beszéd esetében találtunk csak nagyobb szórást (három érdemjegynyi különbséget az ítéletekben). A két legjobban beszélő mondatainak minősítése jónak mondható, hiszen a kísérletben nem szakemberek vettek részt. A kísérlet második részének, az érthetőségre vonatkozó adatoknak az összesítését a 4. táblázat tartalmazza.

A két kísérlet adatait összevetve azt látjuk, hogy a beszédminősítésre adott ítélet rosszabb, mint az érthetőségre adott. A kettő közötti eltérés pontosan jelzi azt, hogy az átlagos hallgató számára a nyelőcsőbeszéd nem megszokott; zörejessége, szaggatottsága, egyéb akusztikai eltérései következtében nem tekinthető nagyon jónak.

Az adatok azonban arra is utalnak, hogy a laryngectomizáltak beszéde elfogadható; a beszéd 70%-osan jó minősítése csak a 75%-nál gyengébb érthetőség (azaz a 4-esnél gyengébbnek ítélt érthetőség) esetén csökken jelentősen. A minősítés és az érthetőség között óhatatlanul

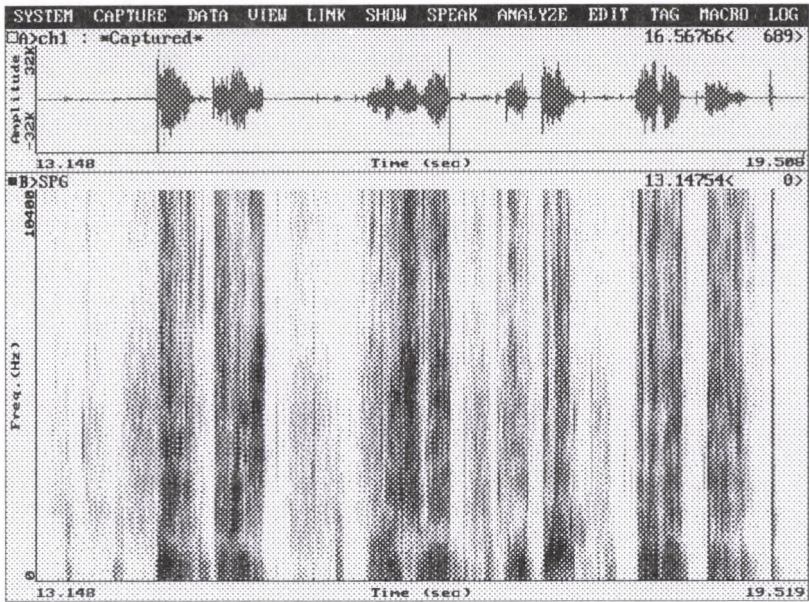
szoros a kapcsolat; a szakember szerint közel azonosan beszélő laryngectomizáltak szubjektív megítélésében a nem szakember hallgatók sem mutatnak lényeges eltéréseket. Ezen a módon két nagyobb csoportra oszthatók a kísérletben részt vevő betegek: az 1., 2. és 3. sorszámúak a "jó beszélők", míg az 5., 6., 7. és 8. sorszámúak a gyenge beszélők. (A 4. sorszámú megítélése közepesnek tekinthető.) Egyetlen olyan beszélőnk volt, akinél a beszédérthetőséget gyengébbnek találták a beszédminőséghez képest (8. beszélő). A két érték közötti különbségnek azonban nincs jelentősége. Megnéztük, mi az összefüggés a szubjektív minősítés és az objektív akusztikai paraméterek között egy a) "jó" és b) egy "gyenge" beszélőnél (vö. 1. és 8. beszélő).

4. táblázat: A laryngectomizáltak beszédének érthetősége

Beszélők	Szubjektív ítélet (%)		
	átlag	szórás	érdemjegyközelítés
1. beszélő	97,6	75-100	5
2. beszélő	97,6	75-100	5
3. beszélő	96,1	75-100	5
4. beszélő	76,3	50-100	4
5. beszélő	52,8	50- 75	3
6. beszélő	52,8	50- 75	3
7. beszélő	34,7	25- 50	2-3
8. beszélő	25,5	0- 50	2

a) A beszédminősítés eredménye: 75,2%, a beszéd szubjektív érthetősége 97,6% (érdemjegyekkel kifejezve: 5-ös az érthetőség és 4-es a minőség). A kialakított nyelőcsőhang mint gerjesztés akusztikailag részben periodikus, azaz kvázi-periodikus szakaszok találhatók benne. A 100-500 Hz-es tartományban 1-2 intenzív felhang található. A magánhangzók formánsszerkezete általában két intenzív formánst mutat (ritkábban egy és három is előfordul). A résmássalhangzóknál 1 vagy 2 intenzív zörejjóc elkülöníthető. Az átlagintenzitás burkológörbéje - az adott mondattól függően - csökkenő vagy domború lefutású. A zörejelemek intenzitása a beszéd egészét tekintve átlagosan 20 dB-lel

alacsonyabbak, mint a jelintenzitás. A beszédtempó 7-10 hang/s, vagyis közelíti/eléri a normál, átlagosnál kissé lassabb beszéd sebességét (1. ábra).

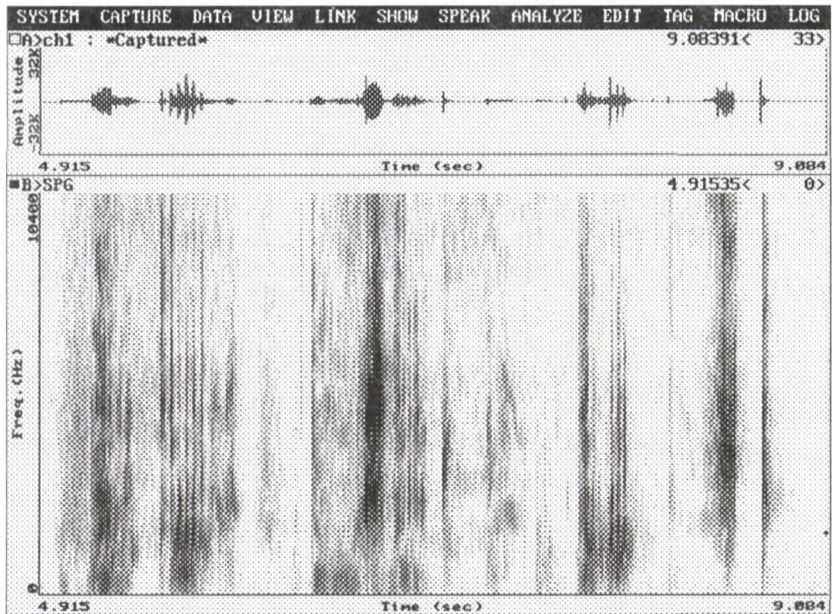


1. ábra

Jól beszélő beteg beszédének hangszínképe (az ábrázolt mondat:
Húsvét Pünkösöd után következik.)

b) A beszédminősítés eredménye: 29,5%, a beszéd szubjektív érthetősége: 25,5% (érdemjeggyel kifejezve: mind az érthetőség, mind a minőség 2-es). A kialakított nyelőcsőhang mint gerjesztés akusztikailag aperiodikus, egyetlen mondatban sem lehetett kvázi-periodikus szakaszokat találni. A 100-500 Hz-es tartományban 1-1 intenzívebb felhang látszott, ezek intenzitása azonban nem sokkal haladta meg a spektrumot elfedő zörejnyalábok intenzitását. A jelintenzitás és a fedő zörejnyalábok intenzitáskülönbsége átlagosan 15 dB. A beszédhangok akusztikai szerkezete nehezen volt elkülöníthető a gerjesztő hang zörejeitől, amelyek részben folyamatosan, részben szakaszosan,

nagyjából azonos intenzitással lefedték a teljes beszédartományt. A magánhangzóknál általában egyetlen formáns volt elkülöníthető biztonsággal. A résmássalhangzóknál egy intenzívebb zörejjóc megjelent, de nem függetlenül az adott kontextustól. Az átlagintenzitás burkológörbéje kivétel nélkül homorú lefutású, a kezdet és a vég között jelentősen mértékben változó intenzitáscsúcsokkal. A beszédtempó lassú, a normál köznyelvi beszéd sebességét nem közelíti meg, az átlagérték: 6 hang/s (2. ábra).



2. ábra

Gyengén beszélő beteg beszédének hangszinképe (az ábrázolt mondat: *Húsvét Pünkösöd után következik.*)

Az elemzett akusztikai adatok és a percepció kísérlet eredményei jól korrelálnak egymással (3. ábra).

gerjesztés

periodikus-----részben periodikus-----aperiodikus

felhangtartomány

100-500 Hz: in-
tenzív elemek-----1-2 intenzív elem-----nincs intenzív elem

időtartam

12 hang/s-----8 hang/s-----6 hang/s

invariancia

formánsok
mássalhang-
zó zörejek-----F1 és F2, 1-2-----1-1 formáns, zörejsáv
zörejjóc

intenzitás

ereszkedő
burkológörbe-----domború burkoló-----homorú burkoló-
görbe görbe

érthetőség

jó-----elfogadható-----gyenge

3. ábra

Az objektív akusztikai adatok és a szubjektív ítélet összefüggései

Következtetések

Az elmondottak alapján következtetéseinket pontokba foglalva összegezzük. 1. Minden laryngectomizált beteget érdemes megtanítani a

nyelőcsőhang-beszédre; a hallgató szubjektív értékelésének eredményei ezt megerősítik. A beszédérthetőségre kapott eredmények ugyancsak arra utalnak, hogy az átlagos hallgató feldolgozási folyamatai ezt lehetővé teszik a nyelvcsőhang-gerjesztésű beszéd dekódolását. 2. Az emberi beszéd oly mértékben redundáns, hogy a tökéletlenebb akusztikai szerkezetű beszéd is érthető marad. 3. A tumor kiterjedése erősen befolyásolja a későbbi beszédet, mert a heges özofágusz-bemenet nehezebb nyelést és eructatiót eredményez, így a beszéd szaggatottságát idézi elő. Még az ennek következtében gyengébben beszélő betegek beszéde is alkalmas a társadalmi érintkezésre. 4. Eredményeink klinikai alkalmazást is lehetővé tesznek, amely segítséget ad a foniáternek, hogy saját szubjektív véleményét objektív adatokkal egészítse ki. Egyetlen - megfelelően kialakított - mondat bemondatása és rövid ideig tartó akusztikai elemzése alapján objektíven is minősíthető a kialakított beszéd (vö. 1. ábra). Az ábra akár klinikai alkalmazásra is megfelelő, hiszen az adott beteg beszéde jól jellemezhető a két végpont között.

Irodalom

Beckmann, G.: *Phonetische Untersuchungen an Laryngektomierten*. HNO 9. 1954, 9.

Frint Tibor: *A hangképzés és zavarai, beszédzavarok*. Medicina. Budapest 1982.

Dénes László: *Az oesophagushang sajátosságairól laryngectomy után*. Orvosi Hetilap 1938.

Kenessey László: *A laryngectomizáltak beszédéről*. Fül-orr-gégegyógyászat 2. 1957, 79-83.

Sáfrán Antal: *A nyelvcsőbeszéd rehabilitációs kérdései*. M.F.F.L.T. Kongresszusa, Referátum. Makó 1981.

Sáfrán Antal: *A beszéd gége nélkül; a nyelvcsőbeszéd akusztikus szerkezete*. Kandidátusi értekezés. Budapest 1988.

Seeman, M.: *Phoniatische Bemerkungen zur Laryngectomy*. Arch. Klin. Chirurg. 140. 1926.

Stern, H.: *Der Mechanismus der Sprech- und Stimmbildung bei Laryngektomierten*. Denker-Kahler Bd. Hdb. 5. 1924.

SZÁMELEMEK KIEJTÉSÉNEK FONETIKAI VIZSGÁLATA

Olaszy Gábor
MTA Nyelvtudományi Intézete

Bevezetés

A beszéd szupraszegmentális elemeinek tanulmányozásával foglalkozó munkákban nemigen találunk olyanokat, amelyek a számok kiejtési szabályaival foglalkoznak. A magyar szakirodalomban egyedül Olaszy közöl olyan adatokat (1989), amelyek a számok kiejtésének szabályait próbálják leírni. A számok fontos elemei a szövegeknek és amennyiben nem egyjegyűek foglalkoznunk kell a kiejtésükre vonatkozó tendenciákkal, szabályszerűségekkal, hogy ezen szövegrészek szupraszegmentális szerkezetét is egyértelműen jellemezni tudjuk. A többjegyű számokat (pl. valaminek az ára, számlának az egyenlege) a szövegben ritkán írjuk le szövegalakban, általában számformájukban szerepelnek. Azt, hogy kiejtésükkor milyen szabályok érvényesülnek ebben a tanulmányban próbáljuk tovább feltárni és rendszerbe foglalni. Az eredményeket az automatikus beszédelőállítás és a beszédfelismerés területén lehet majd használni.

Problémafelvetés

Az eddig kifejtett, számokat is felolvasó rendszereknek az a jellemzője, hogy a számokat a közismert számelemek emberi beszédből kivágott és eltárolt változatainak összekapcsolásával állítják elő a kimondáshoz. A számok automatikus felolvasásához a következő 25 számelemet használhatjuk:

nulla, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, tizen-, 20, huszon-, 30, 40, 50, 60, 70, 80, 90, 100, ezer, millió, milliárd.

Ezekből az elemekből bármilyen szám magyar nyelven összerakható. Ilyen elemekből építkezik néhány bank telefonon felhívható számlaérték felolvasó automatája, néhány telefonszám változást bejelentő automata stb. Meghallgatva ezeket a számbemondó rendszereket feltűnik, hogy azok a számokat szaggatottan, természetellenes prozódiaival, az össze-

kapcsolás határpontjain megjelenő amplitúdó- és dallamív-különbségek miatt "döcögösen" ejtik ki. A hallgató nagyon érzi, hogy a rendszer által összeállított szám a fenti elemekből került generálásra és érzi azt is, hogy ezen automatikusan felolvasott számok kiejtése, prozódiai szerkezete, hangsúlyozása, természetessége messze elmarad az élőbeszédben ejtett számokétól.

Ez annak a következménye, hogy a beszédszolgáltatást megtervező szakemberek nem vették figyelembe a beszédképzés szabályait, a koartikulációból adódó törvényszerűségeket, az ezzel kapcsolatos nyelvészeti és fonetikai tényeket.

Elemzési szempontok

A vizsgálatokban a 25 számelemből indultunk ki. A teljes jellemzéshez három fizikai szinten kell a vizsgálatokat elvégezni, akusztikai (a), intenzitás (i) és időszerekezet (t). Az akusztikai szintet két részre kell bontani, nevezetesen a koartikulációból (k) adódó formánsváltozásokat figyelembe vevő szintre és az alaphangmagasság (h) (az alaphangmagasság és a szószintű elemek) változását leíró szintre. Ebből adódik, hogy ideális esetben minden számelemet 4 paraméteres függvény szerint kellene meghatároznunk, amely függvényben a koartikulációból adódó változatok egy további több paraméteres függvény formájában szerepelnek. Az új elvek alapján tehát egy számelemet a következő általános összefüggés szerint lehet származtatni.

Új számelem = számelem (a (k,h),i,t)

Jelen munkában a fő szempont az időszerekezet, valamint az alaphangmagasság változásának vizsgálata volt. A koartikulációból adódó akusztikai változásokat és az intenzitás szerkezet alakulását Olaszky fent említett munkájában már tárgyaltuk és itt csak érintőlegesen foglalkozunk velük.

Az alábbi kérdésekre kerestük a választ:

1. Milyen a számot felépítő számelemek időtartama a számban való elhelyezkedésük szerint?

2. Hány elemtípust kell megkülönböztetni, ha a számot az emberi ejtéshez közel álló ritmikával akarjuk megvalósítani?

3. Milyen az alaphangmagasság változása a számon belül? Vannak-e hangsúlyozásra utaló alaphangmagasság-kiemelkedések?

4. Milyen az alaphangmagasság változásának általános tendenciája a szám kiejtésénél?

Vizsgálati anyag

A kísérletekhez három felnőtt bemondó ejtésében összesen 720 öt- és hatjegyű számot rögzítettünk hangszalagra, majd ezeket időszerkezet- és alaphangfrekvencia-változás szempontjából analizáltuk.

Módszer

Az időszerkezeti vizsgálatok során megállapítottuk az adott számban szereplő számelemek határait, majd ezután megmértük a számelemek időtartamát. Háromféle számelemet különböztettünk meg a mérési eredmények csoportosításánál: a szám első eleme, belső elem, a szám utolsó eleme. Ezek szerint az 1652844 számból az alábbi elemekre kaptunk adatot:

Kezdő elem: EGY

Belső elem: **millió, hat, száz, ötven, kettő, ezer, nyolc, száz, negyven**

Utolsó elem: négy

A továbbiakban a kezdő elemeket nagybetűvel, a belsőket félkövér kisbetűvel, az utolsókat normál kisbetűvel jelöljük.

A számokat úgy állítottuk össze, hogy minden számelemre minden helyzetben közel egyenlő eloszlásban kapjunk adatot.

Hasonlóan vizsgáltuk az alaphangmagasság-változást is.

Eredmények

Az időszerkezetre vonatkozó mérések átlagolt eredményeit minden számelemre az 1. táblázatban adjuk meg

1. táblázat: A számelemek időtartama

Számelem	kezdő	belső	utolsó
egy	170 ms	240 ms	300 ms
kettő	380	440	620
három	360	380	500
négy	260	320	380
öt	180	220	250
hat	230	270	350
hét	300	320	380
nyolc	350	350	380
kilenc	400	430	510
tíz	240	280	300
tizen....	350	420	---
húsz	340	410	450
huszon...	340	290	---
harminc	450	450	480
negyven	430	440	480
ötven	350	365	380
hatvan	400	380	420
hetven	430	450	480
nyolcvan	460	460	480
kilencven	520	580	610
száz	210	250	380
ezer	280	300	340
millió	-	500	580
milliárd	-	540	600

A táblázat eredményeihez a következő megjegyzéseket fűzzük. Mint látható, a kezdő elemek a legrövidebbek, a záró elemek pedig a leg-hosszabbak. Az eltérés mértéke az adott elemtől függ. Néhány helyen eltérés tapasztalható, például a *huszon...* tekintetében, ahol a közbenső elem rövidebb, mint a kezdő. A *nyolc* elemnél pedig nem tapasztaltunk eltérést a kezdő és a belső elem időtartam között. A fentiekből követ-

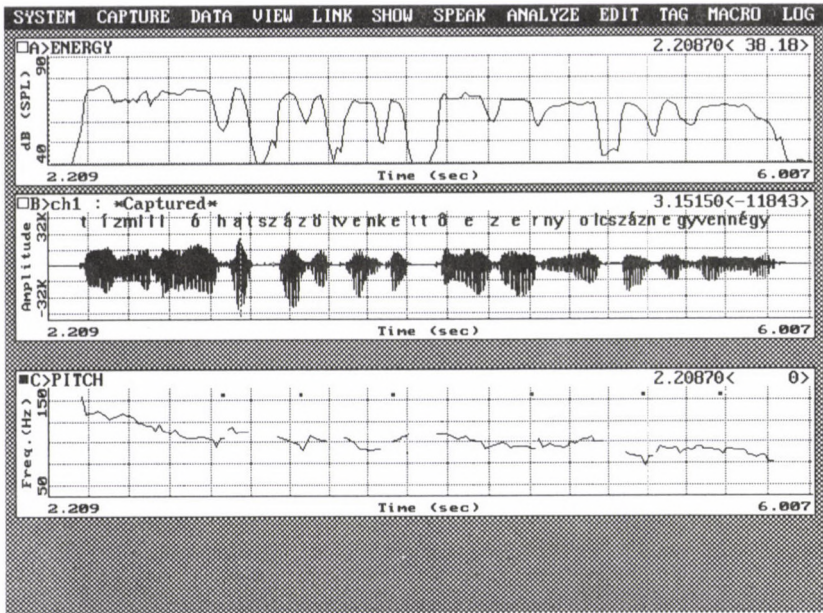
kezik, hogy minden számelemet legalább háromféle időtartammal kell megvalósítani, hogy közelítsük a természetes ejtés ritmikai szerkezetét.

Az alapfrekvencia tekintetében a következő általános tendenciákat állapítottuk meg. A számok kiejtése során az alapfrekvenciagörbe követi a közlés alapgörbéjének vonulatát (l. Olaszy, ebben a kötetben). Erre az alapfrekvencia-görbére szuperponálódnak a számelemen belüli hangsúlyozást megvalósító emelkedő-eső formátumok (l. ábra). Ezek mértéke általában +15% az alapgörbéhez képest. A számok most kimutatott belső hangsúlyozási szerkezetéből arra következtetünk, hogy a szám dekódolásakor és kiejtésekor minden számot úgynevezett számelem-szavakra bontunk és így ejtjük ki őket. Például a 10652844 számot a következő hét számelem-szóval ejtjük ki:

1652844 = tízmillió, hatszáz, ötven, kettőezer, nyolcszáz, negyven, négy

Ezekben a számelem-szavakban az első szótagi hangsúlyozási szabály működik, amint azt az 1. ábra alapfrekvencia (c) görbéjéből le lehet olvasni. Minden számelem-szó önálló eső tendenciájú alapfrekvencia-változással bír, amelynek az indulási frekvenciája mindig magasabban van, mint az előző elem befejezésekor jelenlévő frekvencia. Tehát az alapfrekvencia egy fűrészfogazathoz hasonló görbe szerint változik, amelynek egy általános tendenciája: gyengén eső. A számelem-szavak határait a görbe fölötti kis pontokkal jelöltük meg.

Ha most összehasonlítjuk, hogy milyen elemeket kell használnunk az időszerkezet szempontjából, és milyeneket az alapfrekvencia-változás szempontjából, akkor láthatjuk, hogy igen nagy az átfedés a két elemhalmaz között.



1. ábra

A 10652844 szám kiejtett változatának akusztikai diagramjai

Időszerkezeti elemek:

EGY, millió,hat,száz,ötven,kettő,ezer,nyolc,száz,negyven, négy

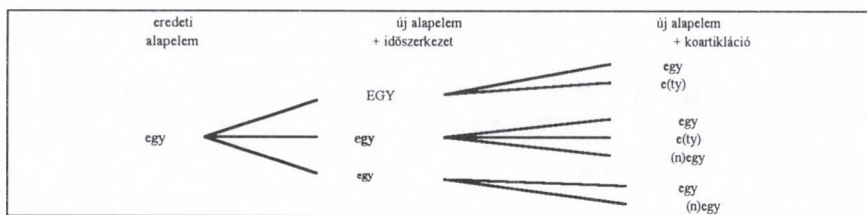
Alapfrekvenciából adódó elemek:

egy*millió* , hatszáz, ötven, kettő*ezer*, nyolc*száz*, negyven, négy

Láthatjuk, hogy a számot leíró elemekben csak a helyiértéket kifejező elemek (*száz*, *ezer*, *millió*) helyzete különbözik. Ezek az elemek mindig hangsúlytalan formát mutatnak, mivel a száelem-szó második, nem hangsúlyos részében szerepelnek. Ebből következik, hogy adott szám korrekt idő- és alapfrekvencia-szerkezet szerinti leírásához az építőelemeket ugyanazon elvek szerint lehet meghatározni és ez az elv megegyezik az időszerkezet szerinti felosztás korábban leírt (kezdő, belső, utolsó) szabályaival. Ha ezek szerint határozzuk meg az építőelemeket,

akkor azok automatikusan tartalmazni fogják a szám belső hangsúlyozását megvalósító alapfrekvenciaváltozásokat is.

A fentiekből következik, hogy ha csak az időszerkezetet és az alapfrekvencia-változásokat szeretnénk korrektül jellemezni szám belső akusztikai szerkezetének leírásakor, akkor $25 \times 3 = 75$ féle számelemet kell megkülönböztetnünk. Ha ehhez hozzávesszük, hogy a számelemek határára fellépő legfontosabb koartikulációs hatások is figyelembe veendők, akkor a 75-ös szám megsokszorozódik. Példaképpen bemutatjuk, hogy az "egy" számelem hány változatban szerepelhet a kiejtés során, ha idő- és alapfrekvenciaszerkezeti, valamint koartikulációs szempontból is korrekt jellemzést akarunk megvalósítani (2. ábra). Az ábrában zárójeles betűvel jelöltük a koartikulációs hatásból eredő hangváltozást.



2. ábra

Az "egy" számelem előfordulási változatai számok kimondásánál

Eredmények

A fenti fonetikai vizsgálatok megmutatták, hogy a számok kiejtésekor milyen szabályok érvényesülnek. E szabályok alkalmazásával egyrészt meg lehet tervezni egy adott szám kimondásához tartozó építőelemeket, illetve e szabályok értelmezésével kiegészítő információkat lehet adni beszédfelismerőknek, amelyek alapján pontosabb döntéseket hozhatnak a felismerés folyamatában.

Irodalom

Olaszy Gábor: Elektronikus beszédelőállítás. Műszaki Könyvkiadó, Budapest 1989.

Ez a kutatás az OTKA TO13565 támogatásával készült.

A LEXIKÁLIS HOZZÁFÉRÉS ÉP ÉS ZAVART FOLYAMATAI

5-10 éves gyermekek szóaktivizálási képességei

Szalai Enikő

MTA Nyelvtudományi Intézete

Bevezetés

A lélektani és a gyermeknyelvi kutatások régóta evidenciaként kezelik, hogy a mindenkori nyelvtudás megítélése és értékelése nem függetleníthető az aktuális korosztálytól. Abban a kérdésben azonban kevesebb az egyetértés, hogy mely tények, vonások és összefüggések a legalkalmasabbak a vizsgált korosztály nyelvi tudásának jellemzésére a nyelvhasználat széles és változatos gyakorlatából. Történetileg az egyén mentális lexikona, illetve a lexikont felépítő egységek struktúrális elemzése került a vizsgálatok középpontjába.

A mentális lexikon felépítésére vonatkozó eddigi ismeretek alapján, a vizsgálati eljárások az aktív és a passzív összetevő elkülönítésével már kísérletet tettek a lexikon két - nem feltétlenül egyforma nagyságú, egymással az aktuális beszédhelyzetekben eltérő mértékben kapcsolatban álló - komponensének differenciált megismerésére. Egyetlen vizsgálati módszer sem tűzi ki célul azonban a mentális lexikon harmadik összetevőjének, az ún. *aktivizálható szókincs*nek a vizsgálatát, annak ellenére, hogy a kommunikációs helyzet sikerét legnagyobb mértékben ennek az összetevőnek a hatékony használata biztosítja. Az aktivizálható szókincs meghatározására Gósy (1993) a következő definíciót használja: "Azon szavak összessége, amelyek a beszédpercepció folyamatokban a kiváltó (akusztikai) inger hatására extra energiával töltődnek fel, azaz aktivizálódnak", tehát az asszociációs folyamatok révén a szükséges egységek "előhívhatókká" válnak mind az aktív, mind pedig a passzív szókincsből.

Felmerül a kérdés, hogy a lexikon elemeinek a fenti hármas (aktív-passzív-aktivizálható) szempont alapján történő csoportosításán túl, az elemek milyen belső szabályrendszer alapján struktúrázódnak, illetve milyen összeköttetések létesülnek az egyes elemek között. A kutatók

között nincs egyetértés arra vonatkozólag, hogy az egységek szemantikai mezőkbe (pl.: *fut, szalad, rohan, lohol* szavak; *kellemes - kellemetlen* szemantikai ellentétpárok), fonetikai relációkba (pl.: *sűn, sūt, sül, süllyed, süveg...*) csoportosulva, avagy grammatikai/szintaktikai relációk formájába (pl.: *megyek, mész, megy, megyünk ... a megy* szócsoporthoz, ill. külön a *megy* fogalomkör és a hozzá kapcsolódó grammatikai rendszer paraméterei) rendeződve építik fel a mentális lexikont. Vitatott kérdés tehát az is, hogy az alapelemek (szavak) szótári alakok, tehát izolált szavak formájában avagy grammatikai jelekkel ellátott formában tárolódnak-e a rendszerben. Az eddig használt lexéma, zárt és szabad morféma, illetve fonéma kategóriák a mentális lexikonban milyen értelmezést nyerhetnek? E kérdéskör megválaszolása különösen az agglutináló nyelvek esetében létfontosságú, ugyanis nem zárható ki az a feltételezés, amely szerint a mentális lexikon szerveződése nem univerzális, hanem erősen nyelvspecifikus (vö. Gósy 1993).

A szókészlet struktúráinak feltérképezése azonban még mindig nem elegendő annak a folyamatnak az ismeretéhez, amely során a szót reprezentáló akusztikai jelsorozat és a mentális lexikon adott egysége közötti megfeleltetés (azaz szóértés) létrejön. A folyamatot, amely során - mint hallgató - a beszédpercepció fázisain keresztül eljutunk a keresett szóig, illetőleg - mint beszélő - előhívjuk ezt a beszédprodukciónál, *lexikális hozzáférésnek* nevezzük. Ha sikerül megértenünk, hogyan történik a lexikális hozzáférés folyamata, hogyan választódnak ki a szavak a lexikonból, közelebb juthatunk a beszédprodukciónál és a beszédpercepciónál egyéb kérdéseinek megválaszolásához is.

A lexikális hozzáférés fázisai

A *beszédprodukciónál* során végbemenő, a lexikális hozzáférési folyamatokra vonatkozó elméletek megegyeznek abban, hogy a folyamat két szakaszra bontható:

1. A kívánt lexikai egység (az ún. *lemma*) kiválasztása a lehetséges alternatívák ezreiből, tehát a *lexikális szelekció*, valamint

2. a kiválasztott egység fonetikai formájának a lexikonban tárolt fonológiai reprezentánsával (az ún. *lexeme*) való összevetése, azaz *fonológiai kódolása*.

Levett (1993) elgondolása szerint a közlésre szánt üzenet a lexikális szelekciót megelőzően egy olyan pre-lexikális struktúrálódáson megy keresztül, amely egy lexikális koncepciót juttat az ún. *formulátorba*, amelynek a feladata az üzenet gondolati formából nyelvi formába való átalakítása. A formulátor kétszakaszos működése során elsőként a (1.) a lexikális szelekció megy végbe, amelynek során az input (gondolati tartalom) aktiválja a mentális lexikon egy olyan elemét, amely megfelel az üzenet szemantikai feltételeinek. Ez egy "felszíni" struktúrát közvetít a következő fázisba, ahol végbemegy a (2.) fonológiai kódolás. A lexikális szelekció során létrejövő "felszíni" struktúrák (a lemmák) szemantikailag és szintaktikailag specifikusak, fonológiailag azonban nem.

A kiválasztott lexikai elemnek a beszédprodukciónak szükséges, megfelelő fonológiai programnak való megfeleltetése történik a fonológiai kódolás fázisában, amelynek végeredményeképpen a formulátort fonetikai paraméterekkel felruházott elemek hagyják el. A szavak fonológiai mintái nem előre elkészített, bármikor felhasználható sablonok formájában tárolódnak, hanem - amint a nyelvbotlások kutatása egyértelműen bebizonyította - újra és újra konstruálódnak az aktuális beszédhelyzetnek megfelelően. Az olyan hibák, mint például *rejttem a fejtvényt* (*fejtem a rejtvényt* helyett) bizonyíthatja, hogy a szavaknak legfeljebb csak a "váza" lehet előre meghatározva, függetlenül attól, hogy milyen szegmensek "töltik ki" a keretet a kódolás során (ld. bővebben Shattuck-Hufnagel (1979) "slot-and-filler" elméletét).

Míg a lexikális hozzáférés modelljeinek nagy része két szakaszra bontja a folyamatot, számos kutató foglal állást amellett, hogy nincs létjogosultsága a két szakaszra tagolásnak, hiszen a két folyamat egymástól nem függetleníthető, átfedések vannak a működésük során, amelyek korrigálhatják (vagy éppen ronthatják) a másik folyamat hibáit (eredményeit). Ezt támasztja alá, hogy a fonológiai kódolásban jelentkező zavarok nagyobb gyakorisággal eredményeznek értelmes szavakat, mind azt a véletlen esélyek alapján elvárhatnánk

(Baars–Motley–MacKay 1975). (Magyar nyelvű anyagon végzett percepció kísérletről ld. Gósy 1992, 88).

A lexikális hozzáférést a *beszédpercepció* során reprezentáló modellek közül kiemelkedik **Morton logogen teóriája** (1969). Elmélete az *aktivációs*nak nevezett *modellek* körébe tartozik, mivel véleménye szerint a mentális lexikon ún. *logogenek* gyűjteménye, amely egységek mindegyike rendelkezik az összehasonlításhoz szükséges jegyekkel és stratégiákkal. A beszédpercepció során az aktuális beszédhelyzetnek megfelelő, szemantikailag aktív kognitív rendszer szolgáltatja a releváns információkat a logogenek számára. Az adott szemantikai információ egy meghatározott aktivációs szintet elérő impulzussal aktiválja a számára megfelelő logogent, így ez extra energiával töltődik fel, és saját szavának fonológiai kódját küldi az ún. "válasz-tárba", ahonnan az artikulációs válasz elindulhat.

Míg tehát az aktivációs modellekben a szavak önmagukat hasonlítják a bejövő jelhez, az ún. *lexikális kereső modellekben* a sorozatos összehasonlítások eredményei egy központi *komparátorban* összegződnek, ez veti össze a bejövő jelet a szavak egy meghatározott csoportjával. A lexikális kereső modellek közül kiemelkedik **Marslen-Wilson cohort elmélete** (1989), amely azt állítja, hogy a szöfelismerés folyamatában valamennyi azonosan kezdődő szó aktivizálódik az adott akusztikai input (cohort) következtében. Ha a cohort ('sereg') egyetlen szóra redukálódott, akkor megtörtént a felismerés, akkor is, ha a szó egészében nem hangzott el. E feltételezés alapja, hogy minden szónak létezik egy ún. felismerési pontja, amelytől kezdődően a szó mással nem cserélhető fel, azaz biztonsággal bejósolható. Az elmélet felhasználja azt a tézist, miszerint az elhangzó beszédet *Markov-láncként* kell felfognunk. A Markov-lánc lényege, hogy a meglévő egységekből nagy valószínűséggel bejósolható a következő egység (Gósy 1992).

Tekintettel arra, hogy magyar nyelvű anyagon még csak elenyésző számú kísérlet témája (Gósy 1993) volt a szöfelismerés folyamatának vizsgálata, jelen kísérletsorozatunk célja, hogy az eddig még csak a felnőttekkel végzett kísérletekre alapozott hipotézisek ismeretében, gyermekpopulációra vonatkozó adatokat, ismereteket adjon a mentális lexikon működéséről, a lexikális hozzáférés sajátosságairól.

Figyelembe véve, hogy a szófelismerés, szóértés folyamata alapvetően a beszédpercepciós mechanizmusok által determinált; hipotézisünk szerint a szóaktivizálás képességének fejlettsége és a beszédészlelési szintek fejlettsége egymással összefüggést mutat, így a zavart lexikális folyamat feltételezi a gyenge beszédpercepciós teljesítményt is. Ha a hipotézis beigazolódná, akkor kijelenthetjük, hogy a szókincsaktivizálási gyengeség a beszédészlelési zavarok egyik mutatója lehet.

A beszédmegértési folyamat első szakaszán túl - amelyben a beérkező ingerek akusztikai, fonetikai, fonológiai szintű feldolgozása zajlik -, a jól működő mentális lexikon, a nyelvi jelek szemantikájának pontos ismerete a beszédmegértés folyamatának második szakaszában, a nyelvi kódrendszer értelmezése során különösen fontos szerephez jut. E logikai menet alapján a szóaktivizálási folyamat eredményessége tehát nemcsak a beszédmegértési folyamat alsóbb, az észlelésben szerepet játszó szintjeinek működésével, hanem a felsőbb, a szemantikai dekodolást végző szintjeivel is - ideértve még az asszociációs szintet is - összefüggésben kell, hogy álljon. Mindkét hipotézisre választ kerestünk a vizsgálat során.

Anyag és módszer

A vizsgálatban 208 éphalló, ép intellektusú, ép beszédprodukción, de gyenge beszédpercepciós, beszédmegértési készségekkel rendelkező gyermek, valamint - kontrollcsoportként - 91 éphalló, ép intellektusú és beszédprodukción és ép beszédpercepciós, beszédmegértési készségekkel rendelkező gyermek vett részt. A gyermekeket 5 életkori csoportba osztottuk (vö. 1. táblázat).

A kísérlet során a GMP-Beszédészlelést és beszédmegértést vizsgáló tesztcsomag (Gósy 1989) egyik altesztjét (GMP11) választottuk a vizsgálat eszközeként, amelyben a gyermekek feladata az volt, hogy a *ma*- és a *ke*- hívószótag segítségével mondjanak szavakat. A vizsgálat időkorlátja 8 s volt, tehát a szógyűjtést addig folytattuk, amíg 8 másodperc szünet nem következett az utolsóként elhangzó szó után.

Figyelembe véve, hogy a szerialitás felismerési és reprodukálási képessége a legfontosabb tényezők egyike a mentális lexikon bővítése szempontjából, a beszédészlelési teljesítmény és a szókincsaktivizálási

képesség összevetéséhez a beszédészlelési teljesítményre vonatkozó adatokat a gyermekek GMP-tesztcsomag 10-es altesztjében produkált eredményei szolgáltatták.

1. táblázat: A kísérletben részt vevő gyermekek életkori megoszlása

Életkori csoportok (év; hó)	kísérleti csoport	kontroll csoport
1. csoport: 5;0-5;11 évesek	41 fő	18 fő
2. csoport: 6;0-6;11 évesek	44 fő	17 fő
3. csoport: 7;0-7;11 évesek	41 fő	20 fő
4. csoport: 8;0-8;11 évesek	43 fő	15 fő
5. csoport: 9;0-9;11 évesek	39 fő	21 fő
Összesen:	208 fő	91 fő

Az alteszt 10 értelmetlen hangsor (logatom) visszamondatásával vizsgálja a beszédészlelési és a beszédprodukciós rendszer összefüggését, célzottan a szeriális észlelés fejlettségének meghatározására irányulva. A tesztben szereplő két, három és négy szótagú logatomok szerkezete megfelel a magyar fonotaktikai sajátosságoknak. A gyermekeknek a tesztben szereplő értelmetlen hangsorokat (*galalajka, zseréb, trankün, sísizdami, feréndekek, bakógy, menelékej, jacolov, vucsityó, kriszposztyüvan*) vizuális segítség - az ajakartikuláció látványa - nélkül kell visszamondaniuk.

A beszédmegértési folyamat és a szóaktivizálási képesség kapcsolatának elemzéséhez a beszédmegértési szintet egy speciális, a mondatértés vizsgálatára kifejlesztett teszt segítségével vizsgáltuk (GMP16). A feladatban szereplő 10 mondat szintaktikai, morfológiai és szófaji szerkezetét tekintve a hatodik életévre - normál beszédfejlődési ütemet feltételezve - elvárható nyelvi szintnek felel meg. A teszt bináris döntéshelyezetre épül, azaz minden tesztmondathoz egy az eredetivel szemantikailag ellentétes mondat tartozik, amelyben a szemantikai különbséget egyetlen, vizuális formában is megragadható részlet

tartalmazza. A teszt összeállításának kritériumai: figyelembe veszi (i) a rövid idejű verbális memória működését, (ii) hogy létezzen a mondatok szemantikai ellentettje, (iii) mindkét mondat egyszerűen ábrázolható legyen, (iv) egyaránt tartalmazzon nyelvtanilag egyszerű és összetett formákat is.

Egy-egy mondatpár jelentését tükröző két színes képet egyszerre mutattunk a gyermekeknek, majd a tesztmondat elhangzása után a mondatot ábrázoló kép kiválasztására kértük őket.

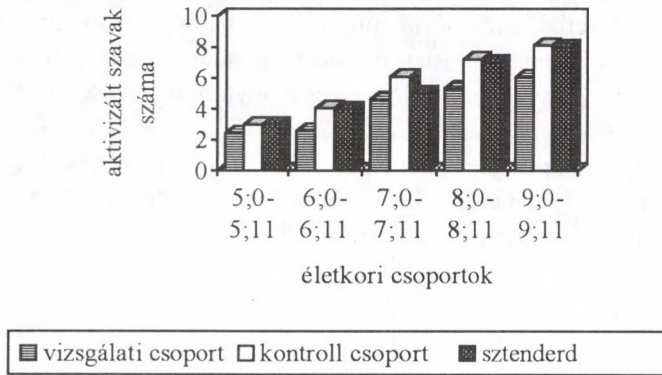
A tesztmondatok:

1. *Az egérke majdnem eléri a sajtot.*
2. *A maci és a nyuszi fára másztak, és az egyikük leesett.*
3. *A macska az asztal mögül húzza az egeret.*
4. *A kislánynak oda kell adnia a könyvet a kislfiúnak.*
5. *Mielőtt a maci ivott, evett egy kicsit.*
6. *Az asztalról leeső gyertyáról beszélnek.*
7. *Nem a nyuszi vette fel a kockás nadrágot.*
8. *A medve szalad, nehogy megcsípjék a méhek.*
9. *Mivel nagyon esett a hó, a kislány mégse ment el szánkózni.*
10. *A kislány megette volna a tortát, ha elérte volna a tálat.*

A vizsgálati eredmények statisztikai feldolgozását a Microsoft Excel 4.0 programmal végeztük.

Eredmények

I. A feldolgozás során elsőként a vizsgálati személyek szókincsaktivizálási tesztben elért eredményeit viszonyítottuk a GMP-tesztcsomag által megadott sztenderd értékekhez (vö. Gósy 1995, 26). Az 1. ábra az aktivált szavak számát és az átlagértékek viszonyát szemlélteti. Az eredmények alapján jól látható, hogy a gyenge beszédészlelési, -megértési teljesítményt mutató gyermekcsoportoknál egyik életkorban sem érte el a vizsgált gyermekek teljesítménye az átlagértékeket, tehát a szókincsaktivizálási képességük átlagosan gyengébb volt az életkoruknak megfelelő szintnél. A kontrollesoport eredményei megfelelnek a sztenderd értékeknek.



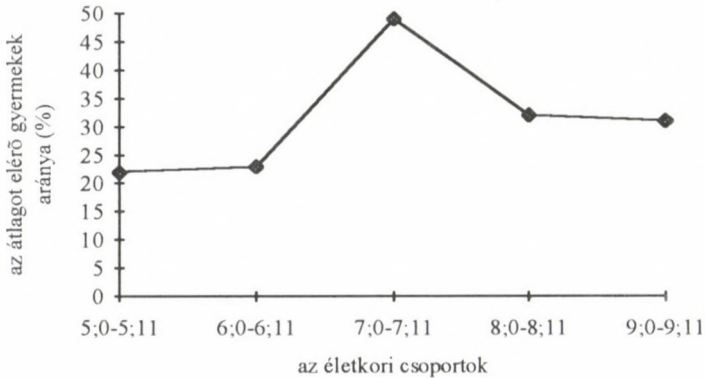
1. ábra

Az aktivizált szavak száma a kísérleti és kontrollescsoportnál életkori bontásban és a sztenderd értékek

Az átlagadatok elemzése nem nyújt elegendő információt arra vonatkozólag, hogy milyen eloszlást mutat a gyermekek teljesítménye a szóaktivizálási feladaton belül, éppen ezért az adatok elemzése során kitértünk annak vizsgálatára, hogy az átlagérték-mutatókon túl, a vizsgált populációban a gyermekek hány százaléka teljesített a normál övezeten belül. Az átlagértéket elérő vagy annál jobban teljesítők arányának változását a 2. ábra szemlélteti.

Az átlagosan vagy átlagosnál jobban teljesítő gyermekek aránya a hetedik életévig lineárisan emelkedik, majd ekkor ugrásszerűen megnő, amit az iskolakezdés kedvező hatásával magyarázhatunk. A mentális érési folyamatok, az ismeretek bővülése, az aktív tanulás, a növekvő számú kommunikációs helyzetben való részvétel pozitív hatásait feltételezve a teljesítmény lineáris emelkedésének folytatódnia kellene, azonban az adatok szerint a fejlődés nemcsak, hogy megtorpan, de a teljesítmény vissza is esik nyolc- és kilencéves korra. A minta kis száma miatt egyértelmű magyarázatot nem adhatunk a jelenségre. Lehetséges magyarázatként azonban felvetődhet, hogy az életkor előrehaladásával, az osztályfokok, az iskolai követelmények emelkedésével a beszédészlelési, beszédmegértési készségekben való elmaradás

egyre nagyobb mértékben befolyásolja kedvezőtlenül a nyelvi, így az iskolai teljesítményeket is.



2. ábra

A sztenderd értéket elérő vagy azt meghaladó gyermekek arányának változása

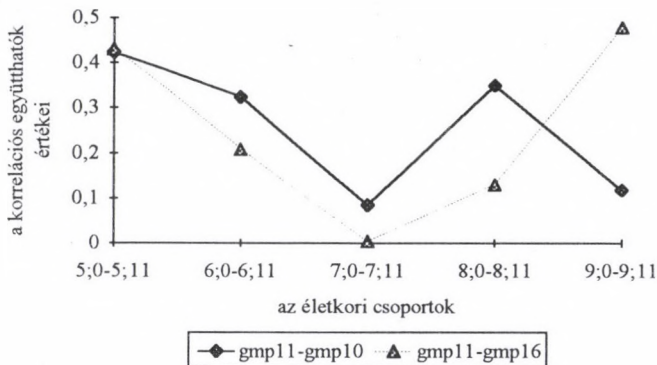
Amint a normál övezeten belül teljesítő gyermekek számának elemzése is előrevetíti, a beszédpercepció folyamat alsóbb és felsőbb szintjeinek tökéletes működése a hatékony szóaktivizálás elengedhetetlen feltételei, tehát a gyenge beszédpercepció teljesítmény kihat az e feladatban mutatott teljesítményre. A hipotézis bizonyítására a gyermekek beszédészlelési folyamatának működéséről felvilágosítást nyújtó tesztben (GMP10) elért eredményeit vetettük össze a szóaktivizálási eredményekkel.

A 208 fős mintán végzett matematikai számítások szerint a **beszédészlelési és a szóaktivizálási teljesítmény között van összefüggés**, a korrelációs együttható értéke $r=0,3623$, tehát leíró matematikai statisztikai értelemben a változók között közepes korreláció áll fenn ($0,3 < r < 0,7$).

A szóaktivizálási teljesítmény és a beszédmegértés kapcsolatát vizsgálva e két változó között is közepes korrelációt találunk

($r=0,3643$), tehát elmondhatjuk, hogy e két nyelvi teljesítmény is **összefügg egymással**.

A mintát életkori csoportokra bontva a korrelációs együtthatók értéke jelentős eltéréseket mutat, ezáltal az vizsgált populációra megállapított összefüggést nem terjeszthetjük ki minden egyes korcsoportra. A szóaktivizálási és a beszédészlelési tesztben nyújtott teljesítmények korrelációs értéke 7 éves korig együtt változik, folyamatosan csökken, majd a 7 éves kortól - az iskolába lépés időpontjától - a korreláció változása ellentétes tendenciát mutat. A 4. és 5. korcsoportban kapott korrelációs együtthatók váratlan eltérését okozhatja a kis elemszámú minta, valamint az a tény, hogy az 1. és 2. korcsoportban vizsgált gyermekek homogén populációt alkotnak (mindannyian közép-, ill. nagycsoportos óvodások), míg a 3. korcsoporttól kezdve a életkor nem határozza meg egyértelműen a vizsgált gyermekek iskolai osztályfokát. Így figyelembe kell vennünk, hogy azonos életkori csoportba sorolt gyermekek akár három különböző osztályfokot is képviselhetnek, ami a teljesítmények szempontjából erősen befolyásoló tényező. A korrelációs együtthatók változásait az 3. ábra szemlélteti.



3. ábra

A korrelációs együtthatók értéke az egyes életkori csoportokban

Az adatok eddigi elemzése során kitűnt, hogy a 7 éves kori iskolakezdés tájékán a szóaktivizálásban átlagosan vagy annál jobban teljesítő gyermekek számának addigi lineáris emelkedése (vö. 2. ábra) – avagy a beszédészlelési-szóaktivizálási és beszédmegértési-szóaktivizálási teljesítmények közötti korreláció lineáris csökkenése – megtörik. Az éles tendencia-változást meghatározó tényezőket próbáltuk felderíteni a vizsgálatban felhasznált tesztfeladatokban elért eredmények szórásának (vö. 2. táblázat) és átlagának (vö. 3. táblázat) elemzésével.

A szóaktivizálásban nyújtott teljesítmények szórása egyenletes emelkedést mutat 7 éves korig, a 7 és 8 éves kor között meredekebb változás áll be, majd 8 éves kortól az eredmények stagnálnak.

2. táblázat: Az egyes tesztfeladatok eredményeinek szórása

Életkori csoportok (év; hó)	GMP11	GMP10	GMP16
5;0-5;11	1,32	20,52	16,43
6;0-6;11	1,66	22,99	14,50
7;0-7;11	1,91	21,99	13,49
8;0-8;11	2,73	20,85	10,52
9;0-9;11	2,63	21,17	10,40

A szórások F-próbával történt összehasonlítása során a változásokat $p=0,05$ szinten szignifikánsnak találtuk. A 7 éves kort megelőző két életkori csoporthoz viszonyított nagyobb fokú szórás-változásra a 7 és 8 éves kor között magyarázatként felhozhatjuk az iskolai oktatás szókincsbővítő hatását, illetve azt a tényt sem hagyhatjuk figyelmen kívül, hogy a vizsgálatok tanúbizonysága szerint (vö. Nagy 1980, 61) a családi nevelés gyakorolja a legnagyobb hatást a gyermekek szókincsére, az óvoda elhanyagolhatóan csekély mértékben fejleszti a nyelvi bázist, tehát az első osztályba eltérő szocio-kulturális környezetből érkező gyermekek nagymértékben különböznek a szókincs tekintetében. Ha elfogadjuk ezt a gondolatmenetet, akkor pozitív hatásként kellene értékelnünk azt, hogy – az eredmények alapján – az iskola az óvodával ellentétben nagyobb mértékben fejleszti a gyermekek

szókincsét. (Ezt támasztja alá az aktivált szavak átlagának szignifikáns emelkedése is.) Az életkori változások nyomonkövetésével azonban látszik, hogy az iskola a fejlesztő hatása mellett tökéletesen megfelel "homogenizálási" feladatának, hiszen az életkor előrehaladtával az egyéni különbségek fokozatosan csökkenek. (Félő, hogy a folyamat nem kizárólag a gyengébben teljesítők átlaghoz való felzárkóztatásának irányában működik.)

3. táblázat: Az egyes tesztfeladatok átlageredményei

Életkori csoportok (év; hó)	GMP11 (db)	GMP10 (%)	GMP16 (%)
5;0-5;11	2,41	54,14	75,60
6;0-6;11	2,56	60,45	76,13
7;0-7;11	4,65	70	83,17
8;0-8;11	5,34	72,44	88,13
9;0-9;11	6,07	72,3	85,64

A 2. táblázat adatai alapján látható, hogy a beszédészlelési teljesítmény (GMP10) szórása az életkorral fokozatosan csökken 6 éves kortól, míg az átlag fokozatosan nő egészen 8 éves korig, majd stagnál (3. táblázat). Az empirikus szakirodalmi adatok szerint az anyanyelvi percepció bázis megszilárdulása körülbelül 6-10 éves korra tehető (vö. Gósy 1992, 133), amely magyarázza a szórások csökkenő, valamint az átlagértékek emelkedő tendenciáját. Az átlagadatokban 8 éves kortól bekövetkező 80%-os teljesítményszinten való állandósulás figyelemfelhívó bizonyítéka annak, hogy a beszédpercepció készségek ontogenetikus fejlődése egy bizonyos életszakaszban leáll, függetlenül attól, hogy a készség elérte-e a megfelelő fejlettségi szintet avagy sem; tehát a beszédpercepció készségek fejlődésében a természetes érési folyamat nem egyedüli meghatározó tényező, a zavart fejlődési folyamat kompenzálására ilyen esetekben ún. "külső" segítségre, speciális terápiára van szükség.

II. A szóaktivizálási tesztben (GMP11) kapott eredmények minőségi szempontú elemzése során 1. a mentális lexikon felépítésére, 2. a lexikális hozzáférés folyamatainak szabályszerűségeire, 3. a zavart folyamati működés eredményeire, illetve okaira vonatkozó megállapítások megfogalmazását tűztük ki célul. A feldolgozás e fázisában a kontrollcsoport eredményeit is felhasználtuk.

A vizsgálati metódus a két különböző szótag segítségével tesz kísérletet az aktivizálható szókincs elérésére, kiküszöbölve ezzel a vizsgálat objektivitását korlátozó ama tényezőjét, hogy eltérő szókezdetekre eltérő az aktivizálható szavak mennyisége és minősége. Mindkét hívószótag esetében a jelen kísérlet eredményei is - hasonlóan az egyéb szókincsvizsgálatok eredményeihez - szófaji eloszlás tekintetében **főnévi dominanciát** (78,6%) alapítottak meg az összes vizsgált gyermekcsoportban. A második leggyakoribb szófaj az ige (13,4%), majd ezt a melléknév (6,7%) követi. A határozószavak és névmások aránya elenyésző (1,3 %).

A szófaji bontáson belül a grammatikailag jelölt, illetve jelöletlen szavak arányát vizsgálva azt tapasztaljuk, hogy a főnevek esetében elenyésző százalékban kaptunk toldalékolt szavakat, tehát a főnevek nagy valószínűséggel **szótári alakok formájában tárolódnak a mentális lexikonban**. Az igék esetében egyes szám harmadik személyű igealakok - tehát szokásosan szótári alakok - aktivizálódtak túlsúlyban, de itt gyakrabban jelent meg a toldalékolt forma (pl.: *marad, maradtak, maradsz, maradj*).

A kísérleti személyek válaszainak összesítéséből kitűnik, hogy milyen módon változnak a lexikális hozzáférés során az egyes szavakban az ún. felismerési pontok (vö. Marslen-Wilson cohort elmélete), amelytől számítva egy és csakis az adott szó érhető el a beszédmegértés, ill. a beszédprodukciónál. A vizsgálatban aktivizált szavak balról jobbra haladó tengely mentén való felismerési útját elemezve megállapíthatjuk, hogy az előhívott szavak többsége az ún. második szinten helyezkedik el, tehát a hívószótag záró magánhangzóját követően egyetlen mássalhangzó után már differenciálódik a szavak nagy többsége.

A vizsgálat tapasztalati megfigyelése, hogy a gyermekeknél a **szóaktivizálás nem lineárisan**, a fokozatos differenciálódást követve **megy végbe** (pl.: *mag*→*maga*→*magas*), hanem random sorrendben történik (pl.: *mama*, *mag*, *magyar*, *Marika*, *macska*).

A súlyos beszédpercepciószavarokkal küzdő gyermekek esetében tapasztalhattuk az aktivizálás folyamatának "megakadását", amikor is a gyermek egy adott szónál "megtapad", és a célszó toldalékolt formáinak ismételtetésével (pl.: *mama*, *mamám*, *mamája*) a legszűkebb hatókörben maradva próbál a vizsgálati szituációnak megfelelni.

A nem a feladatban megjelölt hívószótaggal kezdődő értelmes szavakat, illetve a kívánt hangsorral kezdődő, de értelmetlen hangsorokat soroltuk a tévesztések csoportjába (vö. 4. táblázat). A tévesztések arányainak (1:8 a vizsgálati csoportban és 1:13 a kontroll csoportban) különbsége is alátámasztja, hogy a beszédészlelés szintjeinek – ebben az esetben különösen a fonetikai és fonológiai szintek – zavarát, a hangdifferenciálási képesség fejletlenségét a szókincsaktivizálás során.

A mentális lexikon megfelelő egységének elérési kudarcai jól tükrözik a hozzáférés mechanizmusának lépéseit, a folyamat sikeres kivitelezéséhez szükséges alfolyamatokat. A szeriális észlelés zavara eredményezi a metatéziseket; a fonetikai, fonológiai szintű kódolás, a hangdifferenciálás zavarai pedig a magán- és mássalhangzócsereket.

4. táblázat: Példák a tévesztések típusaiból

<i>A tévesztések típusai</i>	<i>ma-</i>	<i>ke-</i>
1. metatézis	<i>alma</i> , <i>máma</i>	<i>nekem</i> , <i>lepke</i> , <i>gyerek</i>
2. zöngés-zöngétlen csere		<i>gerbera</i> , <i>gcreblye</i>
3. magánhangzó csere	<i>mosás</i> , <i>munka</i>	<i>karom</i> , <i>kanál</i> , <i>Kína</i>
4. mássalhangzó csere	<i>bab</i> , <i>narancs</i>	<i>Teri</i> , <i>pchely</i>
5. értelmetlen hangsor	<i>mandrár</i> , <i>magnák</i>	<i>kedi</i> , <i>kama</i>
6. értelmes szó torzítása	<i>marancs</i> , <i>mangola</i>	<i>kel mel</i>
a hívószótagnak megfelelően (Angola)		(Pall Mall)
7. szemantikai tévesztés	<i>papa</i> (mama)	<i>szerda</i> (kedd)

A mássalhangzócserek esetében a mássalhangzó egy képzési jegyben mindig megegyezik az eredeti mássalhangzóval (pl.: *m-b* csere esetén a képzés helye változatlan, *m-n*, *k-t*, *k-p* csere esetén a képzés módja), tehát itt már a keresési folyamat elindulásakor történik egy hibás szelekció.

Az értelmetlen hangsorok előhívásakor sem beszélhetünk az aktivizálási folyamat teljes kudarcáról, hiszen a hangsorkezdő mássalhangzó, legtöbbször az azt követő magánhangzó is megfelel az aktiválási követelményeknek, valamilyen oknál fogva azonban az aktivációs folyamat "zsákutcába" fut, nem ér el jelentéssel bíró egységet, és a fonotaktikai szabályoknak ugyan megfelelő, de szemantikailag értékelhetetlen hangsort eredményez a beszédprodukción.

Az értelmes szavaknak a hívószótag kritériumainak megfelelő formába történő átalakítása során a lexikális hozzáférés folyamatában már elért, de az aktiválás feltételeinek nem megfelelő egységet egy visszafelé ható torzítással alakítja a folyamat a kívánalmaknak megfelelővé. Az ilyen torzításoknak "engedő" szavak feltételezhetően gyenge, meg nem szilárdult, "felszínesen felvillanó" elemei a lexikonnak, amelyek nem rendelkeznek még az elérés "kiépített útvonalaival", a lexikon többi eleme közötti biztos kapcsolatrendszerrel. (Pl.: az *Angola* szó egy 7 éves gyermek passzív szókincsének is nehezen feltételezhető eleme, ugyanígy a *Pall Mall* márkanév is, amit a reklám – szótanításra egyébként kiváló – ismétlődő frázisai "zúditottak" korrekt fonetikai/fonológiai reprezentáció nélkül a gyermek mentális lexikonába.)

A szemantikai tévesztések minden esetben már egy fonetikailag adekvát egység előhívása után következnek be (*mama*→*papa*, *kedd*→*szerda*). Feltételezhető, hogy a lexikon elemei **többdimenziós kapcsolatrendszerrel** rendelkeznek, amelyek közül csak egy az a fonetikai-fonológiai kapcsolatrendszer, amelyben a keresést elindult. Egy hibás átkapcsolás eredményeként - dimenzióváltással - a fonetikai-fonológiai dimenzióban való továbbhaladás helyett a szemantikai rendszerben halad tovább a keresés.

Következtetések

1. A kísérleti eredmények alapján a felállított hipotézisre, amely szerint (i) a szóaktivizálási teljesítmény és a beszédészlelési teljesítmény, (ii) a szóaktivizálási teljesítmény és a beszédmegértés között összefüggés van, megerősítést nyertünk. Az egyes változók közötti közepes korreláció bizonyítja, hogy a (i) **a szerialitás felismerési és reprodukálási képessége meghatározó, de nem kizárólagos tényező** a mentális lexikon bővítése, adekvát használata szempontjából. A beszédészlelési folyamat egyéb szintjeinek - az akusztikai, fonetikai, fonológiai szint - működése is befolyásolja a lexikális hozzáférés folyamatát; (ii) a tökéletes beszédmegértéshez sem elegendő a nagy elemszámú, jól strukturált mentális lexikon megléte a **grammatikai, szintaktikai relációk** közötti jó tájékozódás képessége is szükséges az eredményes dekódoláshoz.

Mindezek ellenére elmondhatjuk, hogy a **gyermekek szóaktivizálásban nyújtott gyenge teljesítménye megbízhatóan jelzi** a beszédprodukciónál egyértelműen nem diagnosztizálható **beszédpercepciószavart is**.

2. A vizsgált nyelvi készségek területén az életkor változásával nemcsak mennyiségi, hanem minőségi ugrásokat láttunk a fejlődésben. A **7 éves kor** tájára, az iskolakezdés időpontjára tehető az a határvonal, amely a gyermekek teljesítményében éles **minőségi váltást** jelent. A szóaktivizálásban mutatott egyéni teljesítmények közötti különbségek is ebben az életkorban a legnagyobbak, amely különbségeket az iskola az intézményesített oktatás keretein belül egyre csökkenti.

3. Az oktatási folyamatba bekerülő gyermekek több, mint a fele (51%), nem rendelkezik megfelelő szóaktivizálási képességgel, a mentális lexikonban való tájékozódási készsége bizonytalan, ami a nyelvi egységek dekódolása és produkciója során - mindkét funkció adekvát működése szükséges a sikeres olvasás-írás elsajátításához - egyaránt nehézségekhez vezet. Tekintettel arra, hogy a jelen olvasás-írás tanítási módszerek jó része a gyermekek szótalálási tudására épít - tehát olyan készség meglétét feltételezi, amellyel a gyermekek fele nem rendelkezik -, ezen gyermekeket a szótagolás fázisát kihagyó olvasás-tanítási módszerek a szavak balról jobbra haladó "bejöslására", kita-

lálásra készítetik, eleve **kudarcra ítélve ezzel az írott jelrendszer örömteli, problémamentes elsajátításának lehetőségét.**

4. A vizsgálat eredményei alapján a mentális lexikon elementáris egységeinek a szavak **szótári alakját** tekinthetjük. Az eredmények elemzése megerősíti azt a tényt, hogy a **gyermeki szókincs uralkodó szófaja a főnév.**

A vizsgálat során számos olyan probléma merült fel, amelyek megválaszolását a vizsgált minta kis elemszáma, valamint a vizsgálat nem longitudinális volta miatt korlátozott hatáskörűnek kell tekintenünk. E tény magyarázza, hogy arra a kérdésre, hogy az egyes változók között milyen további összefüggések állhatnak fenn, illetve milyen a jelen kísérletben nem vizsgált tényezők (pl.: a rövid távú verbális emlékezet) hatásával kell számolnunk a nyelvi kompetenciák megítélésében, nagyobb elemszámú mintán végzett, még differenciáltabb, longitudinális vizsgálat sorozat adhat választ.

Irodalom

Baars, B. - Motley, M.T. - Mackay, D.: Output editing for lexical status from artificially elicited slips of tongue. In: Journal of Verbal Learning and Verbal Behaviour 14. 1975, 382-391.

Gósy Mária: GMP - Beszédszlelést és beszédmegértést vizsgáló tesztcsomag. Logo-Press. Budapest 1989.

Gósy Mária: A beszédszlelés és a beszédmegértés folyamata. Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola. Budapest 1992.

Gósy Mária: A lexikális hozzáférésről. In: Beszédkutatás 93. MTA Nyelvtudományi Intézete. Budapest 1993, 14-32.

Gósy Mária: GMP-Diagnosztika. A beszédszlelés és a beszédmegértés folyamatának vizsgálata. Nikol. Budapest 1995.

Levelt, W.J.M.: Accesing words in speech production: Stages, processes and representations. In: Levelt, W.J.M.(ed): Lexical Access in Speech Production. Blackwell. Cambridge Massachusetts, 1993, 1-22.

Marslen-Wilson, W.: Access and integration: Projecting sound onto meaning. In: Lexical Representation and Process. Cambridge, Massachusetts, London, England 1989, 3-25.

Morton, J.: The interaction of information in word recognition. In: Psychological Review 76. 1969, 165-178.

Nagy J. József: 5-6 éves gyermekeink iskolakészültsége. Akadémiai Kiadó. Budapest 1980.

Shattuck-Hufnagel, S.: Speech errors as evidence for a serial order mechanism in sentence production. In: Cooper, W.E.-Walker, E.C.T.(eds.): Sentence processing: Psycholinguistic studies presented to Merrill Garrett. Erlbaum. Hillsdale NJ. 1979, 295-342.

A tanulmány a **T 013863** számú OTKA-pályázat támogatásával készült.

SZABÁLYRENDSZER PROZÓDIAI ELEMEEK GÉPI MEGVALÓSÍTÁSÁHOZ

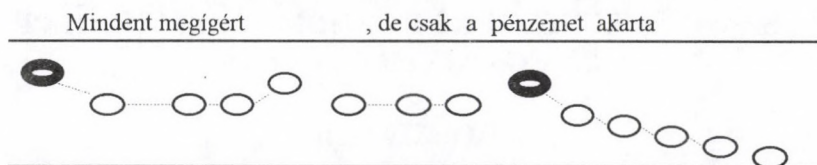
Olaszy Gábor
MTA Nyelvtudományi Intézet

Bevezetés

Az elhangzó beszéd szupraszegmentális elemeinek modellezése bonyolult feladat. Modellezésen itt azt értjük, hogy szabályokat fogalmunk meg, és ezek alkalmazásával megpróbáljuk gépi előállításban megközelíteni az emberi beszéd dallamát, ritmusát, hangsúlyozási szerkezetét. Általános megoldás nem is létezik, hiszen az elhangzó beszéd szupraszegmentális szerkezete sok olyan tényezőtől is függ, amelyek nem modellezhetők. Ilyenek például az adott témakör, amiről elhangzik a megnyilatkozás, a témához kapcsolódó egyéb részletinformációk, előzmények, esetleges jövőbeni tények, amik befolyásolják a beszélőt, a szövegkörnyezeti összefüggések, a beszélő egyéni stílusa, kifejezésformája stb. Mindezek miatt a modellezés csak szűk sávot érinthet. Jelen tanulmányban egy olyan modellrendszert mutatunk be, amivel a magyar és német nyelven elhangzó leggyakoribb közlésformák prozódiai szerkezete írható le paraméterezett formában.

A prozódia megközelítése

A beszéd szupraszegmentális szerkezetének modellezése az utóbbi években került előtérbe. A legtöbb tanulmány az alaphangmagasság változásának szabályaival foglalkozik. A magyar nyelv esetében inkább csak elméleti, leíró jellegű megfogalmazásokat találunk (É. Kiss 1988; Varga 1990, 1994), amelyek sematikus ábrákkal jól érzékeltetik az adott dallamforma vonulatát (1. ábra), de például nem tartalmaznak adatokat a konkrét Hz-értékekre vonatkozóan. Az olvasó számára nem jelent problémát az ilyen ábrák alapján rekonstruálni a kérdéses dallamot, de ezt csak azért tudja megtenni, mert produkciós mechanizmusát a tudatalatti nyelvi, kiejtési szabályok működtetik az imitáció során.

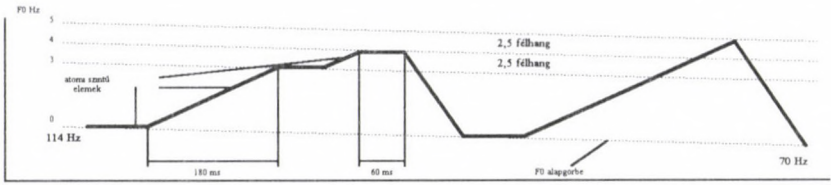


1. ábra

Dallamvonulat ábrázolása Varga módszerével

Ugyanezen dallamot nehéz létrehozni szabályok alkalmazásával, mivel nem tudjuk megadni a szabályok belső adatait (Hz- és időtartamértékek), nem is beszélve arról, hogy a fenti ábránál részletesebb leírásra van szükség ahhoz, hogy ezt a sematizált dallamformát korrekten visszakapjuk gépi megvalósításban. Például meg kell határozni a dallam változásának irányait és mértékét a szótagon belül. Ahhoz viszont, hogy az intonációt felépítő szabályokat megfelelő mélységben meg tudjuk fogalmazni egyrésztől részletesen kell elemezni a beszéd intonációs szerkezetét, másrésztől percepciós tesztekkel kell megállapítani azokat a határvonalakat, amelyek között bizonyos intonációs szabályok működhethetnek. Magyar vonatkozásban ez utóbbira Gósy folytatott kísérleteket (1989) a kijelentő mondatok eső alaphangmagasság görbéje tekintetében, és Gósy–Terken (1994) a magyar eldöntendő kérdés intonációs csúcsának konkrét, percepciósan még elfogadható alsó és felső határértékei vonatkozásában.

Más nyelvekre konkrét kísérleteket is elvégeztek az optimális intonációs modellezés meghatározására. Ilyenek például Collier (1990), Terken és Collier (1990) kísérletei hollandra, Adriaens (1991) munkája németre, Mertens (1995) rendszere franciára. Mindhárom modellben lineáris elemeket alkalmaztak az intonáció leírására, valamint percepciós tesztek is végeztek. Adriaens rendszerében (2. ábra), az intonációs alapgörbe (basic F0 contour) képezi a kiindulási alapot, és erre kerülnek rá a frázis- és szósztintű intonációs változások az idő függvényében. Ez utóbbiakat úgynevezett atomi szintű intonációs elemekből lehet felépíteni egy a kottázáshoz hasonló, 5 szintes rendszerben, ahol szintenként 2,5 félhangnyi távolság van a szintek között.



2. ábra
Adriaens modellje az intonáció felépítésére

Kontrasztív jellegű szabályrendszer leírására Olasz (1991) tett kísérletet, amelyben olyan általános intonációs építőelemeket fogalmazott meg, amelyek felhasználásával több nyelven is fel lehet építeni intonációs görbéket. Hunyadi legutóbbi tanulmányában (1995) érzékelteti azt a hiányt, hogy a hangsúly vizsgálatában leginkább elméleti fejtegetésekre támaszkodunk, de konkrét mérési rendszer nincs kidolgozva.

Kétségtelen, hogy az alaphangmagasság változása a leglátványosabb, és hogy ennek van a legnagyobb szerepe a prozódia kialakításában. Ezt Swerts és Gelujkens kísérletei (1993) is igazolták. E mellett azonban nem elhanyagolható az időtartamok változása, valamint a hangerő változása sem. A természetes beszédben e három paraméterérték kialakítását a biológiai rendszerünk egyszerre végzi el, modellezésnél is erre kell törekedni. Ez azt jelenti, hogy egy adott beszédelem (hang, szótag, szó) prozódiai szerkezetének leírásakor az is előfordulhat, hogy ugyanazon elemben kell alaphangmagasságot, intenzitást és időtartamot módosítani.

Célkitűzés

A kutatás célkitűzését két pontban foglaljuk össze.

1. Olyan modellt kell készíteni, amelyikkel a prozódia három alapvető komponense (alaphangváltozás, hangnyúlás, hangrövidülés és hangerő csökkenés, növekedés) leírható. Az alaphangmagasság-változás tekintetében külön kell tudni kezelni az intonációs alappörbe (base line) meghatározását és külön a hangsúlyozásból, hangsúlytalanításból adódó alaphangmagasság-változásokat. Fontos szempont, hogy az alaphang-

magasság-változások kiszámítása mindig egy adott, előre meghatározott, kiindulási Hz-értékből történjen.

2. A modellezés olyan paramétervektorok sorozatában valósuljon meg, amelyekben mind a paraméterek, mind pedig azok belső értékei egyértelműen, fizikai értékekkel (Hz, dB, időtartamváltozás) vannak leírva, és ez tükrözi az adott szöveg leendő prozódiai szerkezetét.

Elemzés

A prozódiai szerkezet modelljének kialakításához 82 magyar és 32 német mondat kiejtett változatának akusztikai szerkezetét analizáltuk. Az anyagban egyszerű kérdő, kijelentő és óhajtó, valamint összetett kijelentő közlésformák szerepeltek. Az intonációs görbék tipizálása során lineáris közelítésre törekedtünk (hasonlóan Adriaens módszeréhez), és megpróbáltuk szétválasztani az intonációs alapgörbét, valamint az erre szuperponálódott finomabb változásokat. További törekvésünk volt közelíteni a nemzetközi és a magyar intonációs irodalomban is elfogadott emelkedő és eső formákból álló kombinációkat (lásd Vargánál), hogy ezeket tudjuk majd alkalmazni a modell leírásánál.

A vizsgálatok során minden mondatról szinkronizált dallam-, intenzitás- és oszcillogramgörbét készítettünk, majd a regisztrátumokon megjelöltük a beszédhangok helyét is. A munkához a Kay Elemetrics Computerised Speech Laboratory (CSL) számítógépes programot használtuk.

Az elemzés eredményei

A dallamgörbék kiértékelésénél háromfajta változást különböztettünk meg. Az első a dallamgörbe általános vonulata a mondat, vagy frázisegység teljes tartamában. Ezt neveztük alap-dallamgörbének. Frázisegységnek vettük azokat a részeket, amelyekben a dallamgörbe változási tendenciája megfelelt az eső, emelkedő, változatlan (neutrális), emelkedő-eső és eső-emelkedő kategóriáknak. A második változás képezte az alap-dallamgörbére ráépült szótag és szószintű intonációs elemeket (a szóhangsúly különböző fokozatai). A harmadik elemet az első és második egymáshoz viszonyított különbsége alkotta, amikor azt

akartuk megtudni, hogy az alap-dallamgörbe változása milyen hatással van a szószintű változás frekvenciaértékeire.

A hangidőtartam és a hangintenzitás változása tekintetében a korábbi kutatások irányvonalát követtük (Olaszy 1994; 1995), nevezetesen a hangrövidüléseket a többszótagú szavak utolsó magánhangzójában, a hangnyúlást kiemelten hangsúlyozott szavakban, továbbá a mondat utolsó szavában, valamint a hangintenzitás csökkenését a hangsúlytalan szavakban és a mondat végén.

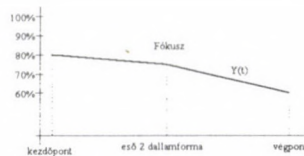
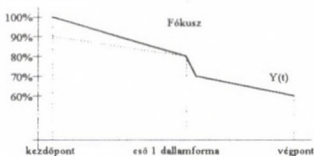
A szabályrendszer

A prozódiai szabályrendszerben 7-féle alap intonációs görbét rendelhetünk a közlésekhez az alábbi sematizált formák szerint. Ezek hordozzák és fejezik ki a mondat vagy frázis alapvető intonációs vonulatát.

Neutrális - összetett mondatokban használjuk. Az alapfrekvencia a 80% környékén mozog.

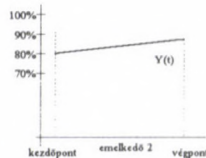
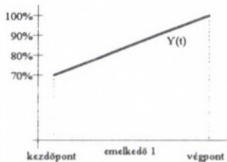
Eső 1 - (erős) amelyet mondatkezdő helyzetben használunk

Eső 2 - (gyenge) amelyet mondat belseji és mondat végi helyzetben használunk

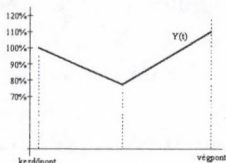
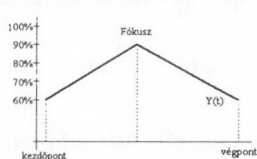


Emelkedő 1 - (erős) főleg rövid, 1, 2 szavas mondatokra használjuk

Emelkedő 2 - (gyengén emelkedő)



Emelkedő-eső és eső-emelkedő



A fenti sematizált alap-dallamformákhoz a következő magyarázatot fűzzük. Az alapelemek kezdő Hz-értékei relatív értékek, amelyeket %-ban adtunk meg. A 100%-os érték a kijelentő mondat görbéjének induló értéke. (Szintézisnél ez felelhet meg az indulási adatként beállított alaphangmagasság-értéknek.) Látható, hogy az egyes dallamformák indulási Hz-értéke különböző %-os magasságokban kezdődik, befejezési értékeikre is ez jellemző. Ezek a különböző értékek valósítják meg az egymáshoz rendelhetőséget. Az egymáshoz rendelési forma azt jelenti egyrésről, hogy a görbék kezdő és befejező Hz-értékei úgy vannak beállítva, hogy azok nem egy kiragadott mondatra vonatkoznak, hanem biztosítják az egymáshoz kapcsolhatóságot, vagyis azt, hogy egy több mondatos közlésegségben az egymás utáni mondatok dallamát törés nélkül, ugyanazon hangfekvésre vonatkoztatva lehet megtervezni. Ez vonatkozik természetesen arra az esetre is, amikor egy mondaton belül akarunk többféle intonációs görbét egymás után kapcsolva megvalósítani. Az $Y(t)$ függvény fejezi ki, hogy az adott görbe az idő függvényében változóként értendő. A fókusz azt fejezi ki, hogy az adott alapgörbét a mondatfókusz (frázisfókusz) megtörheti, két szakaszra bonthatja. Az ábrák értelmezéséből következik, hogy ugrások is lehetnek az egyes intonációs görbék kapcsolódási pontjain, ami megfelel a természetes beszéd egyes speciális intonációs eseteinek is.

Szó- és szótagszintű szabályok

A szavak szintjén mind az alaphangmagasság-változásra, mind a hangidőtartamok rövidülésére, illetve nyúlására, mind pedig az esetleges hangintenzitás változásokra (halkulás, erősödés) meghatároztunk szabályokat. Ezek többnyire a szóhangsúlyozást valósítják meg.

A szó- és szótagszintű intonációs szabályok (építőelemek) "ráülnek", rászuperponálódnak az intonációs alapgörbére. Intonációs szempontból

a következő szókatégoriákat különböztetjük meg: a mondat első szava, utolsó szó a mondatban, mondatfókusz pozícióban lévő szó, fókusz utáni szó, kiemelt és hangsúlyozott szó (több színen), vessző előtti szó, egyéb szó.

Prozódiai szempontól fontos az egyes szavak hosszúsága is. A következő kategóriákat különböztetjük meg: egy szótagú, két szótagú, normál (3-4 szótagú), hosszú (5-6 szótagú), nagyon hosszú (7 szótagtól).

Egy-egy szóra kétfajta intonációs változást írhatunk elő: az egész szóra vonatkozik, illetve a szó valamelyik szótagjára vonatkozik. Ezen intonációs változások megvalósítására a következő intonációs struktúrákat (építőelemeket) alkalmazhatjuk:

A szószintű intonációs változások

I1. Felugrás-- csökkentés

-- általában az első szótagi hangsúlyra kell alkalmazni

I2. emelkedés -- csökkentés

-- általában a szóbelseji hangsúlyra (magyarban az eldöntendő kérdésnél is használható)

I3. emelkedés -- ott maradás

a vessző előtti szó végén és egyéb speciális dallamok előállításánál lehet alkalmazni,

I4. leugrás -- ott maradás

-- általában a hangsúlytalanított szavakra (negatív hangsúly) használjuk

I5. emelkedés, visszaugrás

-- általában a szóvégi hangsúlyra és az egyszótagú szavakra vonatkozik

I6. leugrás--csökkentés

-- általában az utolsó szó kategóriára vonatkozik

I7. leugrás

-- általában az intonációs alapgörbe indulási értékének beállításához használhatjuk

Mint látjuk, a fenti dallamépítő szabályok többsége a szóhangsúly megvalósítását szolgálja. A szóhangsúlyt a következő kategóriákkal valósíthatjuk meg: erős, másodlagos, gyenge hangsúly (ezeket nevezzük

pozitív kategóriáknak). Ezek ellentettje a hangsúlytalan elemek csoportja (negatív kategória), majd végül szükség van egy neutrális elemre, amiben gyakorlatilag nincs intonációs változás, de nem a hangsúlytalan kategóriába tartozik, annál magasabb alaphangmagasság-értéken valósul meg.

A szóhangsúly helye változhat a szón belül, mindig a megjelölt szótagon realizálódik. Ennek megfelelően többfajta építőelemet határoztunk meg a szóhangsúlyozás kifejezésére. Ez a többnyelvűségi alkalmazást teszi lehetővé (például a német nyelvben a hangsúly lehet az adott szó bármelyik szótagján), de a magyarban is alkalmazni kell például a vesszőt kifejező szóvégi intonációs emelkedésnél. Az előbbiből következik az is, hogy ebben a rendszerben a hangsúly terminológia néhány speciális esetben eltér a megszokottól, mivel itt az alaphangmagasság-változása a funkcionális elem, és ennek rendeltük alá az egyes alkalmazási pontokat a hangsorban, nem pedig nyelvészeti kategóriák szerint (lásd, hogy a vessző előtti szó végén az intonációs emelkedést a hangsúlyra kidolgozott elemmel valószínűsítjük meg).

A szószintű intonációs elemek gyakorlati megvalósítása

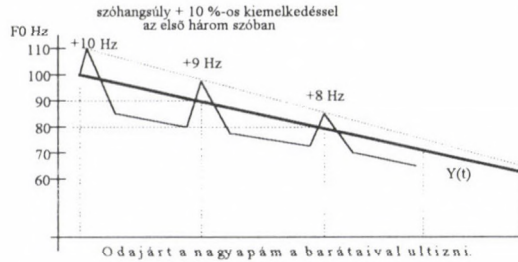
Mint láttuk a szószintű intonációs elemekben (I1...I7) alaphangmagasság-növelést, illetve -csökkentést kell megvalósítani, amit %-ban adunk meg. A változás mértéke az adott szabálytól függően lehet más és más. Ebben a rendszerben ezeket a változásokat az alap intonációs görbére (Y(t)) vetített relatív számítási eljárással számítjuk ki a következő képlet szerint:

$$F0 \text{ tényleges} = F0(Y(t)) \times \text{a megadott } \%$$

A pozitív változások kiszámításához 100-nál nagyobb százalékot, a negatív változásokéhoz 100-nál kisebb százalékot adunk meg. Ezek az értékek nyelvtől függően változhatnak.

Ezzel a számítási formával érjük el azt, hogy az egyes szavakra ráültetett szószintű intonációs struktúrákban a változás mindig más lesz (hiszen az alap intonációs görbe is változik). Eső alapgörbe esetében ez úgy nyilvánul meg, hogy a szóhangsúlyozásból eredő kiemelkedő dal-

lancsúcsok az eső alapgörbe elején lesznek a legnagyobbak, és az ugyanazon hangsúlyhoz tartozóak a végén pedig a legkisebbek (3. ábra). Ez a fokozatos értékcsökkentés módszere, amit az angol szakirodalom "range reduction" néven említ (Collier 1990).



3. ábra

A szóhangsúlyok kiszámításának eredménye: a fokozatos értékcsökkenés

Szabálypéldák

Prozódiai szabályok az első szótagi hangsúly megvalósításához

Dallammozgás: II-es típus: felugrás -- visszacsökkentés

Alapfrekvencia felugrás

Kezdet: az első zöngés hang kezdete

Vége: az első magánhangzó 10%-a

Mértéke:

- erős 1,45 Y Hz
- másodlagos 1,25 Y Hz
- gyenge 1,1 Y Hz
- neutrális Y Hz
- negatív 0,85 Y Hz

Alapfrekvencia (F0) visszacsökkentés

Kezdet: az ugrás vége után azonnal

Vége: -- a szótag magánhangzója utáni közvetlen zöngés hang vége (pl. *mama*)

-- a szótag magánhangzójának vége

(ha nincs utána közvetlenül zöngés hang, pl. *aszta!*)

Végértéke: 0,8 Y Hz

Kivétel: ha a szó első, illetve utolsó szó a mondatban.

Hangidőtartam szabály szóeleji hangsúlyra

a hangsúlyos magánhangzót megnyújtjuk

-- ha rövid magánhangzó, akkor 120%-ra

-- ha hosszú, akkor 130%-ra

Kivétel: a diftongus itt nem számít magánhangzónak, tehát nem kell nyújtani. A többi hang időtartama nem változik

Hangerő szabály a szóeleji hangsúlyra

Hangerő változás nincs a szótagban.

Kivétel, ha fókusz szóról van szó, ekkor a hangerőt növelhetjük 3 dB-lel a szótag magánhangzójában.

A fenti szabályhoz hasonló szerkezetű szabályok írják le a szóhangsúly többi változatát is.

Prozódiai szabályok az egy szótagú szón lévő hangsúly megvalósításához

Dallammozgás: I2-es típus: emelkedés -- visszacsökkentés

Ez a hangsúlyozási forma akkor valósul meg, ha az egyszótagú szó fókuszba kerül, illetve kiemelt szó lesz. Egyébként az egyszótagú szavak a neutrális hangsúlykategóriába sorolandók, kivéve a negatív hangsúlyosakat (névelők, névmások stb.).

Alapfrekvencia-emelés

Kezdeté: az első zöngés hang kezdete

Vége: a magánhangzó 80%-a

Mértéke:

-- erős 1,45 Y Hz

-- másodlagos 1,25 Y Hz

Alapfrekvencia-visszacsökkentés

Kezdeté: az emelkedés vége után azonnal

Vége: -- a következő zöngés hang vége (pl. *én, lány*)

-- a következő szó eleje (ha nincs utána zöngés hang pl. *ott, sok*)

Végértéke: 0,8 Y Hz

Hangidőtartam szabály egy szótagú szó hangsúlyozására: a hangsúlyos magánhangzót megnyújtjuk 150%-ra

Hangerő szabály: a hangerőt növelhetjük 3 dB-lel a szótag magánhangzójában.

Prozódiai szabályok a hangsúlytalan szóra és a mondat utolsó szavára

Dallammozgás: I4-es típus: leugrás -- csökkentés.

Alapfrekvencia leugrás

Helye: a szó kezdete előtt

Mértéke: 0,9 Y Hz

Alapfrekvencia csökkentés

Kezdete: a szó zöngés részének eleje

Vége: -- a szó vége

Végértéke: 0,8Y Hz

Kivétel: Ha a szó fókuszban van, illetve kiemelt szó, akkor ezt a szabályt nem kell végrehajtani, a jelzés szerinti intonációt kell ráültetni.

Kivétel: ha a szó vessző előtt van, akkor a vessző intonációs szabálya érvényesül, az az erősebb.

Hangidőtartam-szabály a hangsúlytalan, illetve az utolsó szóra

Mondatbelseji hangsúlytalan szó hangjait 80-90%-ra rövidíthetjük. Az utolsó szó hangjait 120-130%-ra megnyújthatjuk

Hangerő-szabály: a hangerőt csökkenthetjük 3 dB-lel a szó teljes időtartamára

Eredmények

A fent ismertetett szabályrendszert kísérleti jelleggel alkalmaztuk szintetizált mondatok prozódiai szerkezetének megvalósítására mind magyar, mind német nyelven. A kezdeti meghallgatásos ellenőrzések azt

mutatják, hogy a rendszerrel jól meg lehet közelíteni a természetes prozódiaát. Úgy találjuk, hogy a kialakított prozódiai építőelemekkel a szükséges intonációs formák legtöbbjét meg lehet valósítani. A rendszer alkalmazható összetett mondatokban is. Ilyenkor több intonációs alapelemet kell egymás után kapcsolni, és mindegyikre alkalmazni a szószintű intonációs építőelemek ráültetését. A szabályokban megadott %-os értékek változtathatók, így a rendszer alkalmas a további kísérletezésre, amelynek során a különböző %-os beállítással készített beszédmintákat percepcióos tesztekkel lehet minősíteni, és ily módon még optimálisabb beállítási értékeket meghatározni.

Irodalom

Adriaens, L. M. H.: Ein Modell deutscher Intonation. IPO Eindhoven, 1991.

Collier, R.: Multi-lingual intonation synthesis: Principles and applications. Proc. of the ESCA Workshop on Speech Synthesis. Aufrans (France) 1990, 273-276.

Gósy Mária: Beszédeszlelés. MTA Nyelvtudományi Intézet, Budapest 1989.

Gósy, M.–Terken, J.: Question marking in Hungarian: Timing and height of pitch peaks. Journal of Phonetics 22. 1994, 269--281.

Hunyadi László: Mondathangsúly a magyarban. In: Beszédkutatás 95. Szerk.: Gósy M. MTA Nyelvtudományi Intézet, Budapest 1995, 32-45.

É. Kiss Katalin: Még egyszer a magyar mondat intonációjáról és hangsúlyozásáról. Nyelvtudományi Közlemények 89. 1988, 151--160.

Mertens, P.-d'Alessandro, Ch.: Pitch contour stilization using a tonal perception model. In: COST 233, Prosodics in synthetic speech. Ed. F.J. Lundin, Telia 1995, 82-85.

Olaszy, G.: A crosslinguistic description of intonation contours of a multilingual text-to-speech system. Proc. of the XIIth ICPhS Vol. 4. Aix-en-Provence 1991, 210-213.

Olaszy Gábor: Hangidőtartam-módosító kísérletek a gépi beszéd ritmusának javítására. In: Beszédkutatás 94. Szerk.: Gósy M. MTA Nyelvtudományi Intézet, Budapest 1994, 140-151.

Olaszy Gábor: A kérdés, a figyelmeztetés, a felszólítás és a kérdés prozódiaja a kijelentő mondat tükrében. In: Beszédkutatás 95. Szerk.: Gósy M. MTA Nyelvtudományi Intézet, Budapest 1995, 46-61.

Swerts, M.–Gelujkens, R.: The prosody of information units in spontaneous monologue. *Phonetica* 50. 1993, 189-196.

Terken, J.–Collier, R.: Designing algorithms for intonation in synthetic speech. *Proc. of the ESCA Workshop on Speech Synthesis, Autrans (France) 1990*, 205-208.

Varga László: Stilizált beszéddallamok a magyarban. In: *Élőnyelvi tanulmányok*. Szerk.: Balogh L. és Kontra M. MTA Nyelvtudományi Intézet, Budapest 1990, 169-177.

Varga László: A hanglejtés. In: *Strukturális magyar nyelvtan 2. Fonológia*. Szerk.: Kiefer F. Akadémiai Kiadó, Budapest 1994, 468-546.

Ez a kutatás az OTKA TO13565 támogatásával készült.

